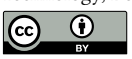
UniqueNFT: Uniqueness Protection of Digital Assets in Decentralized Web

KUN YANG, Artificial Intelligence Research Institute, Shenzhen MSU-BIT University, Shenzhen, China and Guangdong-Hong Kong-Macao Joint Laboratory for Emotion Intelligence and Pervasive Computing, Shenzhen, China
HAIHAN DUAN, Artificial Intelligence Research Institute, Shenzhen MSU-BIT University, Shenzhen, China and Guangdong-Hong Kong-Macao Joint Laboratory for Emotion Intelligence and Pervasive Computing, Shenzhen, China
YUBO ZHAO, Artificial Intelligence Research Institute, Shenzhen MSU-BIT University, Shenzhen, China and Guangdong-Hong Kong-Macao Joint Laboratory for Emotion Intelligence and Pervasive Computing, Shenzhen, China
JIALE CHENG, Artificial Intelligence Research Institute, Shenzhen MSU-BIT University, Shenzhen, China and Guangdong-Hong Kong-Macao Joint Laboratory for Emotion Intelligence and Pervasive Computing, Shenzhen, China
RUNHAO ZENG, Artificial Intelligence Research Institute, Shenzhen MSU-BIT University, Shenzhen, China and Guangdong-Hong Kong-Macao Joint Laboratory for Emotion Intelligence and Pervasive Computing, Shenzhen, China
XIPING HU, Artificial Intelligence Research Institute, Shenzhen MSU-BIT University, Shenzhen, China, Guangdong-Hong Kong-Macao Joint Laboratory for Emotion Intelligence and Pervasive Computing, Shenzhen, China, and School of Medical Technology, Beijing Institute of Technology, Beijing, China

This work was supported in part by the Key Field Projects of Ordinary Universities in Guangdong Province (No. 2025ZDZX3050); in part by the National Natural Science Foundation of China (Grant No. 62576213); in part by the Engineering Technology Research Center for Ordinary Universities in Guangdong Province (No. 2024GCZX005); and in part by Guangdong-Hong Kong-Macao Joint Laboratory for Emotional Intelligence and Pervasive Computing, Artificial Intelligence Research Institute, Shenzhen MSU-BIT University.
Authors' Contact Information: Kun Yang, Artificial Intelligence Research Institute, Shenzhen MSU-BIT University, Shenzhen, Guangdong, China and Guangdong-Hong Kong-Macao Joint Laboratory for Emotion Intelligence and Pervasive Computing, Shenzhen, Guangdong, China; e-mail: jamesykunun@gmail.com; Haihan Duan (corresponding author), Artificial Intelligence Research Institute, Shenzhen MSU-BIT University, Shenzhen, Guangdong, China and Guangdong-Hong Kong-Macao Joint Laboratory for Emotion Intelligence and Pervasive Computing, Shenzhen, Guangdong, China; e-mail: duanhaihan@smbu.edu.cn; Yubo Zhao, Artificial Intelligence Research Institute, Shenzhen MSU-BIT University, Shenzhen, Guangdong, China and Guangdong-Hong Kong-Macao Joint Laboratory for Emotion Intelligence and Pervasive Computing, Shenzhen, Guangdong, China; e-mail: gabrielzhao502@gmail.com; JiaLe Cheng, Artificial Intelligence Research Institute, Shenzhen MSU-BIT University, Shenzhen, Guangdong, China and Guangdong-Hong Kong-Macao Joint Laboratory for Emotion Intelligence and Pervasive Computing, Shenzhen, Guangdong, China; e-mail: chengcharlie052@gmail.com; Runhao Zeng, Artificial Intelligence Research Institute, Shenzhen MSU-BIT University, Shenzhen, Guangdong, China and Guangdong-Hong Kong-Macao Joint Laboratory for Emotion Intelligence and Pervasive Computing, Shenzhen, Guangdong, China; e-mail: runhaozeng.cs@gmail.com; Xiping Hu (corresponding author), Artificial Intelligence Research Institute, Shenzhen MSU-BIT University, Shenzhen, Guangdong, China, Guangdong-Hong Kong-Macao Joint Laboratory for Emotion Intelligence and Pervasive Computing, Shenzhen, Guangdong, China, School of Medical Technology, Beijing Institute of Technology, Beijing, Beijing, China; e-mail: huxp@smbu.edu.cn.



ACM 1559-1131/2026/02-ART17
https://doi.org/10.1145/3779414

With the rapid evolution of the Decentralized Web (DWeb), decentralized technologies have paved new avenues for Web3 applications and the authentication of digital assets. Among them, Non-Fungible Tokens (NFTs) have gained significant popularity due to their immutability and uniqueness, reshaping the landscape of artistic creation, marketing, and intellectual property protection. However, current blockchain-based NFT implementations still face core challenges within decentralized architecture: how to maintain decentralization while ensuring the visual uniqueness of digital assets and reducing storage costs. The rampant issue of duplication undermines the scarcity of digital art and erodes market confidence in copyright authenticity. Moreover, high gas fees and energy consumption further hinder the widespread adoption of NFTs, while reliance on external storage solutions like InterPlanetary File System (IPFS) introduces risks of data instability and loss. To address these challenges, this article presents the UniqueNFT framework, a novel architecture that deeply integrates blockchain oracles with decentralized storage verification mechanisms. The framework achieves three key technological breakthroughs: Using image inversion and generation techniques based on Encoder for Editing (E4E) and StyleGAN3, it extracts compact and expressive semantic features from NFT images, enabling efficient data compression and significantly reducing on-chain storage volume; The Crypto-Mask algorithm, by utilizing the hash value of blockchain user information (user-controlled SHA-256 digest of Ethereum address, user nickname, and registration time), ensures the visual uniqueness of NFTs; A smart contract extension compatible with the ERC721 standard, demonstrating UniqueNFT's seamless integration within the blockchain ecosystem. By leveraging the technologies of the Decentralized Web, our framework represents an important step forward in enhancing the security and uniqueness of digital assets. It not only innovatively resolves the issues of NFT duplication and homogenization but also injects new vitality and long-term momentum into the creation of a trusted, sustainable blockchain-based digital asset ecosystem.

CCS Concepts: • **Security and privacy** → **Privacy protections**; • **Human-centered computing** → *Interactive systems and tools*; • **Computing methodologies** → Image representations;

Additional Key Words and Phrases: Decentralized web, blockchain, non-fungible token, digital assets

ACM Reference Format:
Kun Yang, Haihan Duan, Yubo Zhao, JiaLe Cheng, Runhao Zeng, and Xiping Hu. 2026. UniqueNFT: Uniqueness Protection of Digital Assets in Decentralized Web. *ACM Trans. Web* 20, 1, Article 17 (February 2026), 30 pages. https://doi.org/10.1145/3779414

1 Introduction

Blockchain and Decentralized Web technology have flourished in recent years, giving rise to an unprecedented wave of the digital economy and providing new platforms for the creation, trading, and management of digital assets. In this context, **Non-Fungible Tokens (NFTs)** [84, 88, 90] have rapidly emerged as the key vehicle for representing ownership and uniqueness of digital assets. Their distinctiveness and immutability have brought revolutionary opportunities to art creation, marketing, and intellectual property protection. Classic examples include Chef Sperxos' "Glow in the Dark" NFT, which successfully promoted his unique dining experience in the real world [16], and Dubai's art cars, where limited-edition art cars featuring NFTs have introduced a novel form of car investment [29]. These cases demonstrate how NFTs, powered by Decentralized Web, are creating entirely new digital asset markets [5, 84], improving revenue mechanisms [9] for artists and creators, and significantly enhancing market vitality and consumer engagement.

Although NFTs have demonstrated numerous advantages in promoting digital assets, enhancing artists' compensation, and fostering new business models, several significant issues remain to be addressed. Due to the decentralized nature [6, 14, 93] of blockchain, NFT works can be easily copied, re-minted, and traded on popular NFT trading platforms. This phenomenon has led to the proliferation of unauthorized copies in the market [28, 36, 42], which undermines the uniqueness

and irreplaceability that NFTs should possess, ultimately disrupting the original intent and vision of Web3. In 2017, a virtual cat in CryptoKitties sparked a copyright dispute by using copyrighted images without the permission of the original copyright holder [39]. OpenSea, the world's largest NFT service platform, has also stated that more than 80 percent of the assets created through its simplified "lazy minting" process consist of plagiarized works, fake collections, and spam.[1] This highlights the failure of NFTs to fulfill their promise of protecting digital rights. As this issue of homogenization continues to spread, the problem of replication not only diminishes the rarity of digital art but also indirectly reduces the rightful earnings of artists, potentially undermining the overall health of the NFT ecosystem. Additionally, the high gas fees and energy consumption [50, 51] are significant concerns [24, 54, 87], keeping NFT transaction costs high and limiting widespread adoption. While reliance on external storage systems, such as IPFS [52, 67], alleviates some of the burden of on-chain data storage, it also introduces risks of instability and data loss, further eroding users' trust in the security and long-term value of digital assets.

Motivation. The fundamental concern in the NFT domain is that the promises of uniqueness and immutability are increasingly undermined. First, current architectures rely heavily on off-chain storage systems (e.g., IPFS or centralized servers), which are susceptible to access failures and data loss, undermining the long-term availability of digital assets. Second, current on-chain verification remains limited to basic hash matching, offering little defense against visually similar or semantically duplicated content, which contributes to rampant plagiarism and content homogenization. Third, redundant data storage and high gas fees significantly constrain sustainable on-chain deployment and limit broader participation from creators and users. Therefore, an urgent need for a technical solution that can fundamentally address the NFT ecosystem's challenges in security, persistence, and usability.

Approach. To address the aforementioned challenges, we propose a decentralized NFT storage and verification framework, integrating deep learning, blockchain, and cryptography. By combining image inversion algorithms and AIGC models, we extract semantic features from images to achieve efficient compression and storage of digital assets. The framework utilizes E4E and StyleGAN3 technologies to transform original images into compact latent codes, significantly reducing on-chain data storage volume and thereby substantially decreasing gas fees. Additionally, the system supports personalized editing of image semantic information, ensuring that each NFT possesses unique visual characteristics, truly meeting users' pursuit of "one-of-a-kind" digital assets. To further ensure verification accuracy, the framework authenticates image reconstruction results through blockchain oracles and smart contract hash verification mechanisms, avoiding excessive reliance on traditional IPFS or private servers. Meanwhile, UniqueNFT is designed as a pluggable application-layer framework built atop existing blockchain infrastructure. It follows mainstream NFT protocol standards (such as ERC-721 [63]), enabling seamless integration into various blockchain ecosystems with strong compatibility and scalability.

Contribution. In summary, the main contributions of this article are as follows:

- We propose a novel decentralized NFT storage and verification framework, focusing on enhancing the protection of digital asset uniqueness.
- Through image inversion and reconstruction technologies based on E4E and StyleGAN3, we achieve efficient extraction of image semantic information, effectively reducing storage costs.
- We design a personalized image editing algorithm (Crypto-Mask) that combines blockchain user-information hash, effectively addressing the issue of NFT replication and homogenization.

[1] https://www.theblock.co/linked/132511/opensea-reveals-that-over-80-ofits-free-nft-mints-were-plagiarized-spam-or-fake

2 Preliminary

2.1 NFT Protocols

On blockchain and Ethereum platforms, the commonly used NFT protocols include ERC-721, ERC-1155, and ERC-998. ERC-721 [88] defines the uniqueness of each NFT, making it widely used for trading digital art and collectibles. ERC-1155 [69], on the other hand, enables the management of both fungible and non-fungible tokens within the same contract, making it particularly suitable for batch transactions and virtual items. ERC-998 [55] allows for the combination of multiple NFTs or fungible tokens into a single composite asset, catering to more complex asset management scenarios. These protocols, by providing standardized operational rules, enhance the liquidity of NFTs, reduce transaction costs, and ensure the decentralization and uniqueness of digital assets. These advancements in technology provide a more stable and scalable foundation for the digital asset market.

2.2 Smart Contract

Smart contracts [7, 19, 75] are automated, decentralized protocols that ensure the automatic execution of contract terms in a decentralized network based on pre-set conditions and rules, without the need for third-party intermediaries. This guarantees transparency, verifiability, and efficiency in the execution of transactions and agreements, enabling it widely used in various sectors, including finance [71], supply chain management, and digital assets.

Categorized by functionality, smart contracts range from basic asset transfers to complex **Decentralized Finance** (**DeFi**) protocols. The structure of a smart contract typically consists of three core components [64, 78]: the contract code, the execution environment, and the blockchain network that supports its operation. The contract code defines the conditions and rules for contract execution. The execution environment ensures that these conditions are met and triggers the corresponding actions. The blockchain network guarantees the immutability, transparency, and security of the contract, preventing any unauthorized modifications [60]. Smart contracts have become a fundamental component of blockchain applications. As blockchain adoption continues to grow, smart contracts significantly contribute to the digital economy by minimizing the use of intermediaries and embedding trust directly into protocols, thereby promoting more transparent and efficient economic systems.

2.3 NFT Asset Decentralized Storage

In the domain of NFT asset storage, there are two mainstream models: centralized storage and decentralized storage. Centralized storage depends on controlled servers managed by NFT platforms (e.g., OpenSea,[2] Nifty Gateway,[3] and Rarible[4]) [87] or cloud services (e.g., AWS, Google Cloud), where data governance is monopolized by a single entity. While operational costs are typically borne by platforms (e.g., AWS standard storage costs approximately $0.023 per GB/month [80]), this model harbors fatal flaws: server shutdowns or data tampering can permanently destroy NFT metadata.

Decentralized storage [13, 23] employs blockchain technology to achieve distributed data custody. Representative solutions include **InterPlanetary File System** (**IPFS**) [12] and Arweave [89]: IPFS fragments files into content-addressed encrypted blocks stored across a global node network, with Filecoin's incentive protocol ensuring long-term persistence ($0.00019 to $0.004 GB per month [53]). Arweave adopts a "one-time payment, permanent storage" model, enforcing data immutability via

[2] https://opensea.io/
[3] https://www.niftygateway.com/
[4] https://rarible.com/

cryptoeconomic mechanisms. Core advantages of decentralized storage include single-point failure resistance, authenticity verification, and long-term cost efficiency. These attributes position it as the evolutionary direction for NFT asset storage, exemplified by leading projects like Decentraland migrating fully to IPFS architectures.

3 Related Work

3.1 Digital Image Copyright Protection

Over the past few decades, the protection of digital image assets has been extensively studied, leading to the development of various effective technologies. The most traditional method involves Visible **Watermark** (**WM**) [26], where copyright identifiers are directly overlaid onto digital images, making unauthorized copying and distribution challenging. However, a series of methods for removing visible watermarks [25, 26, 33, 43, 65, 82, 83] have subsequently emerged. As demands for higher image quality and more robust copyright evidence have increased, invisible Digital Watermarking techniques [98] have gained prominence. These methods embed copyright or identity information into the original image in a manner imperceptible to the human eye, ensuring that the watermark can still be accurately extracted after processes like compression, filtering, or geometric transformations. This approach facilitates long-term protection and authentication of images. Digital watermarking techniques are primarily categorized into spatial domain algorithms [18, 27, 44], such as LSB-based methods [15, 57, 85], Patchwork algorithms [96], and transform domain algorithms [98], including those based on singular value decomposition [35, 58] and discrete Fourier transform [59, 76]. These techniques embed copyright information into digital images through specific algorithmic processes, allowing for the recovery of watermark information using predefined keys or reference templates in the event of a copyright dispute.

Blockchain technology [17, 22, 73], with its decentralized, tamper-proof, and traceable characteristics, has emerged as an ideal platform for digital image asset protection. When digital images or NFTs are recorded on a blockchain, every transaction and ownership change is permanently logged and cannot be unilaterally altered. This permanence enables clear tracing of the initial time and owner of the image, providing strong evidence of copyright ownership. Additionally, during the process of minting NFTs, the metadata of digital images is often hashed using algorithms like SHA-256 [68] to generate a unique hash value, ensuring the uniqueness and immutability of the NFT. If unauthorized modifications are made to the image, its hash value will no longer match the original record, thereby detecting tampering.

In recent years, a new paradigm of personalized generation technology based on cryptographic identity markers has emerged in the field of digital rights protection. MetaCube [30] is a decentralized **user-generated content** (**UGC**) editor based on blockchain identity. Its core innovation lies in embedding a user's unique cryptographic credential (such as blockchain users' Ethereum addresses, hash of decentralized web user information or metaverse DID identifier [72]) into the generative neural network training process. In this framework, when users refine low-resolution 3D models into high-fidelity digital assets through a GAN generator, an encryption-driven random dropout mechanism (Crypto-Dropout [31]) is employed. Specifically, this mechanism dynamically controls the random deactivation probability of neuron connections during the forward propagation phase based on the hash value, thereby embedding an irreplicable identity watermark into the generative model's geometric topology and texture details.

Despite notable progress in digital image copyright protection, existing methods still exhibit significant limitations in terms of overall performance and practical applicability (see Table 1). First, visible watermarking embeds explicit identifiers onto image surfaces, offering reasonable anti-copy resistance and a degree of visual uniqueness. However, such overlays often obscure semantic regions [20], degrade visual quality, and are vulnerable to removal via image editing techniques

Table 1. Comparative Analysis of Image Copyright Protection Approaches

Comparison Criteria	Visible WM [26, 65]	Digital WM [18, 27, 59, 76]	NFT Hash ID [13, 17, 22, 73]	Crypto-Dropout [30, 31]	Crypto-Mask (ours)
Visual Uniqueness	✓	×	×	✓	✓
Anti-Copy Resistance	✓	✓	×	✓	✓
Traceability	×	✓	✓	✓	✓
Robustness to Interference	Low	Medium	Strong	Medium	Strong
Computational Overhead	High	Medium	Low	Medium	Medium

(e.g., inpainting), lacking both robustness and traceability. Second, invisible digital watermarking algorithms embed identity information into the underlying features of an image without affecting its appearance, providing a certain level of anti-copy resistance. However, these methods cannot ensure image-level visual uniqueness. Furthermore, since watermark extraction typically depends on external keys or detectors, they lack native traceability and do not support transparent on-chain verification. Third, blockchain hash mechanisms leverage decentralized storage and immutability to provide strong timestamping and traceability guarantees, making them a core component of NFT systems. Yet, as their verification relies solely on static hash values, minor modifications (e.g., resizing, cropping, filtering) can bypass detection, undermining both visual uniqueness and anti-copy resistance. Finally, the Crypto-Dropout mechanism introduces cryptographic identity markers into generative models at the semantic level, offering improvements in uniqueness and anti-copy resistance. However, its reliance on random perturbations during inference often leads to uncontrollable outputs, compromising texture fidelity and edge structure, and thus weakening visual uniqueness and semantic consistency.

In summary, existing approaches fail to simultaneously achieve visual uniqueness, anti-copy resistance, and on-chain traceability in a unified and practical manner. To address these limitations, our proposed UniqueNFT framework builds on blockchain infrastructure to ensure inherent immutability and transparent traceability. Beyond conventional hash- or watermark-based verification methods, we introduce Crypto-Mask, a novel personalization mechanism that modulates semantically meaningful layers within the generative model using hash of user information. In contrast to prior approaches such as Crypto-Dropout, which rely on stochastic neuron deactivation and often compromise image fidelity and semantic structure, Crypto-Mask achieves controlled, deterministic modulation. This allows for the embedding of identity-linked features while maintaining high visual quality. As a result, our framework offers a more robust integration of authenticity, semantic expressiveness, and decentralized verifiability for next-generation NFT copyright protection.

3.2 AIGC-driven Image Editing

AIGC-driven image editing technologies provide dual technical support for achieving NFT uniqueness: generative models construct editable latent spaces, while inversion techniques compress data to fit on-chain storage. This section will explore the technological evolution of each stage in detail.

3.2.1 Disentangled Space in StyleGAN. Understanding and manipulating the latent space of StyleGAN has been a key focus in the research of generative models, as it enables more controlled image generation. Early works by Bau et al. [10, 11] and Shen et al. [74] demonstrated that StyleGAN's [45–47] latent space can be disentangled, allowing for control over specific attributes through linear manipulation in the W space. Further advancements, such as those by Yang et al. [94], showed that the early layers of the network control layout, middle layers handle object presence, and late layers manage fine-grained rendering details. In addition, Studies have shown

that manipulation in W+ space can provide additional flexibility in controlling attributes at different levels of abstraction.

Recent studies have continued to explore the potential of StyleGAN's latent space for various tasks, such as image inpainting, style transfer, and super-resolution [1, 34]. Notably, Nitzan et al. [62] emphasized disentangling identity and pose through fixed decoders, while Richardson et al. [70] explored image translation by encoding sketches into StyleGAN's W space. In our work, we perturb images by editing the S space. According to Wu et al.'s research [91], the S space offers more disentanglement than both the W+ and W spaces. We build upon the work provided by StyleSpace, utilizing the disentanglement operations in the S space to provide finer-grained, attribute-specific controls, thereby enabling richer and more detailed editing.

3.2.2 Image Inversion. The concept of GAN inversion, introduced by Zhu et al. [100]. In this pioneering work, the authors illustrate how performing such an inversion allows for the utilization of the GAN's latent space semantics to facilitate a variety of image manipulation tasks. A number of subsequent studies [1, 2, 21, 47, 79] have approached this problem by directly optimizing the latent vector to minimize the reconstruction error for a given image. These methods generally achieve high-quality reconstructions, but they come at the cost of significant computational time. On the other hand, some methods have proposed the use of an encoder [3, 4, 66, 70, 81, 99] that learns a direct mapping from an image to its corresponding latent vector. While these encoder-based approaches are considerably faster than optimization-based methods, they typically result in lower-quality reconstructions. To address this tradeoff, hybrid approaches have been introduced, using an encoder to initialize the latent vector, followed by further optimization [38, 92].

4 UniqueNFT System

4.1 Architecture and Workflow

At a high level, the architecture of the UniqueNFT system can be divided into an off-chain component and an on-chain component, as shown in Figure 1. The off-chain component primarily consists of two key tasks: the training of an image editing neural network model and the storage and trading of NFTs. The image editing neural network model is divided into two modules: image inversion and image generation. The inversion module employs a latent space encoder [4, 81] to compress input images into implicit feature representations via a mapping network. Utilizing an enhanced W+ space embedding strategy, this module achieves an 8:1 dimensionality reduction rate while preserving semantic fidelity. For image synthesis, the system incorporates the Alias-Free Generative Adversarial Networks (StyleGAN3) [45], which generates identity-unique NFTs by applying cryptographic style vectors derived from hash of user information. On the other hand, the core task of the on-chain component is to adopt the ERC-721 protocol standard to mint compressed and encoded images into NFTs with verifiable uniqueness while managing their full lifecycle through smart contracts on the blockchain. Meanwhile, the system extracts key user information–including Ethereum address, username, and registration time–and uses it as input to the SHA-256 algorithm to compute a user-specific hash. This hash is then utilized to generate a Crypto-Mask, which modulates the decoder to produce a unique style-editing vector. By leveraging the irreversibility and collision resistance of the SHA-256 algorithm, this approach ensures the uniqueness of each NFT's Crypto-Mask and its corresponding perturbation parameters, thereby guaranteeing user-specific scarcity in NFT generation.

To clearly illustrate the execution logic of the framework, the workflow of this system is divided into four stages.

Stage 1: Crypto-Mask. By collecting and utilizing user's information on blockchain, the Crypto-Mask is generated to edit the fully connected layer weights of the StyleGAN3 synthesis

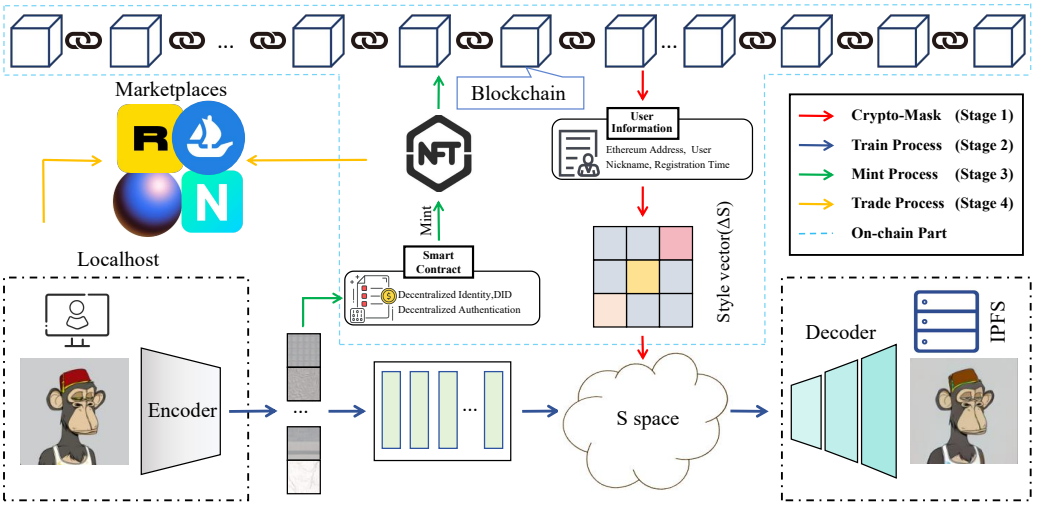


Fig. 1. The overview decentralized architecture of UniqueNFT.

network. Through a bit-to-bit coupling mechanism, each bit of the Crypto-Mask dynamically adjusts the corresponding bit in the weight matrix, generating style control signal that are strictly tied to the blockchain user identity. This signal, after undergoing affine transformation modulation, drives the generated image to exhibit unique features in terms of texture, color, and composition–characteristics that are verifiably linked to the blockchain. Compared to traditional random noise perturbations, this approach ensures the uniqueness and reproducibility of the edit results by leveraging the cryptographic constraints of the hash.

Stage 2: Off-chain Model Training. Both the inversion module (Encoding Module) and the generation module (Decoding Module) are trained concurrently on NFT datasets. They learn the data distribution of the dataset images and align the latent space of the compressed vectors with that of the generation module. A core functionality of the encoder is to achieve efficient image compression: it takes an input image of size 256×256 and encodes it into a latent vector of size 16×512, resulting in a compression ratio of 8:1. After training, the model weights are hashed using the SHA-256 algorithm to generate a 32-byte model fingerprint, which is then written into the ModelRegistry module of the smart contract. The NFTs, with unique user-specific appearance generated through crypto-masking, are stored in a decentralized storage network (such as IPFS, Arweave, Sia, etc.), and the storage credentials are bound to the on-chain metadata via a **Content Identifier (CID)**. The compressed encoding output by the inversion module is recorded in a 256-bit fixed-point format in the smart contract, supporting verifiable inference in subsequent processes.

Stage 3: NFT Minting. This process begins with a hash verification of the personalized generated image: a globally unique asset fingerprint (AssetHash) is created using the SHA-256 algorithm. Next, a metadata package compliant with the ERC-721 extended standard is constructed, containing key attributes such as the decentralized storage credential CID, model version fingerprint, and creation timestamp. This metadata package is submitted to the blockchain network via the MintSemanticNFT function in the smart contract, which includes a built-in verification mechanism to check the uniqueness of the AssetHash and the authenticity of the IPFS/Arweave storage receipt. Upon successful verification, the smart contract generates an encrypted event log containing the complete data of the Crypto-Mask and weight modulation records (employing ElGamal homomorphic encryption) and initializes the ownership address of the NFT to the creator's wallet.

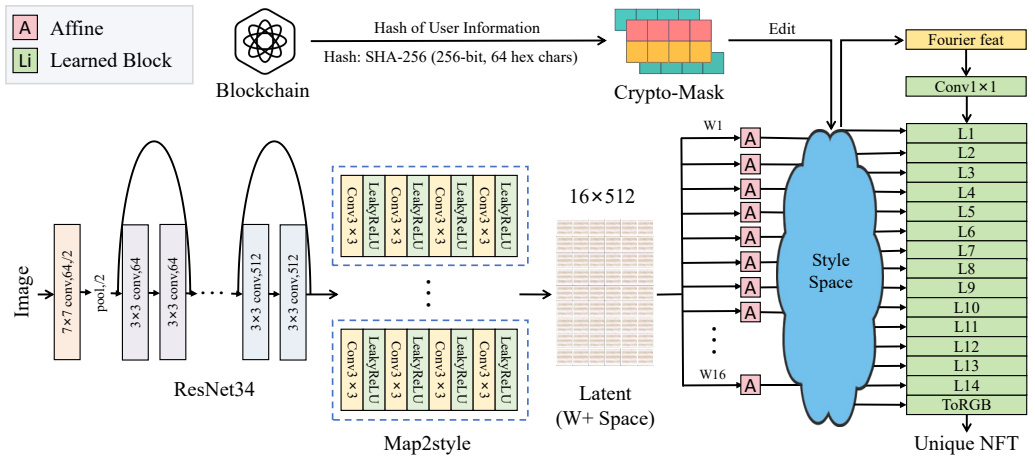


Fig. 2. The neural network architecture in UniqueNFT.

Stage 4: NFT Trading. The transaction process strictly follows the ownership transfer mechanism defined by the ERC-721 standard. When the buyer initiates a transaction request, the smart contract first verifies the caller's management rights to the target NFT by checking the legality of the operation using the isApprovedForAll or getApproved functions. After the permission check is complete, the contract updates the NFT's owner field to the buyer's address and generates a corresponding ownership transfer event on the blockchain. Simultaneously, the smart contract automatically executes tiered royalty distribution: the original creator receives a preset percentage of the share via the EIP-2981 standard interface, the platform service fee is transferred to a multi-signature wallet via an intermediary contract, and the remaining funds are settled to the seller's address via ZK-Rollup batch processing.

4.2 Neural Network Design

The UniqueNFT system adopts a plug-and-play generative architecture, allowing flexible replacement of the intermediate image editing neural network to meet specific requirements. As illustrated in Figure 2, the core neural network consists of two main components: a semantic-aware encoder and a decoder, working together to achieve end-to-end personalized NFT generation. The encoder employs unsupervised learning to extract deep semantic features from NFT images, converting them into compact feature vectors for efficient storage and on-chain verification. The decoder reconstructs images from these feature vectors under frequency-domain constraints. Additionally, during the decoding phase, the system incorporates hash of user information to perform identity-aware editing on specific layers, ensuring the generated NFTs exhibit distinct user-specific styles (further detailed in Section 4.3).

Encoder. The encoder architecture adapts the E4E inversion framework to achieve precise mapping from input images to StyleGAN3's W+ latent space. Utilizing ResNet-34 as the backbone, its cross-layer skip connections effectively capture multi-scale features ranging from local details (textures, edges) to global semantics (pose, illumination), where other feature extraction methodologies are also feasible [95]. The terminal Map2Style module projects fused features into the W+ space through hierarchical fully-connected layers, generating style vectors that satisfy:

$$\min \theta_E \left[\|G(E(x)) - x\| + \lambda L_{LPIPS}(x, G(E(x))) \right], \tag{1}$$

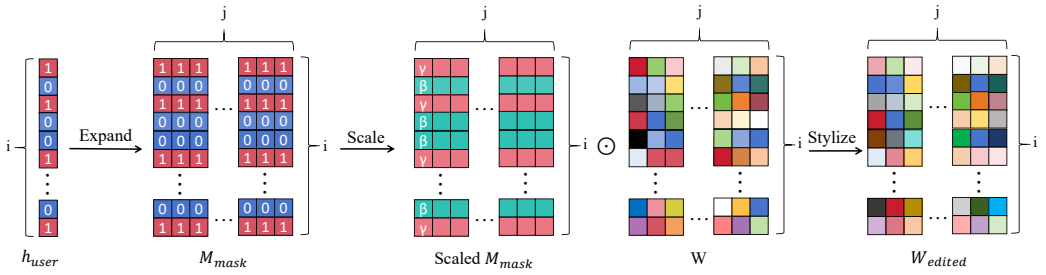


Fig. 3. Process of Crypto-mask algorithm.

where θ_E represents the trainable parameters of encoder E, G denotes the pre-trained StyleGAN3 generator, λ is the weight coefficients for balancing pixel loss and perceptual loss, and L_{LPIPS} corresponds to the perceptual similarity loss based on deep features.

Decoder. The decoder in our study builds upon the StyleGAN3 architecture, which resolves the inherent conflict between geometric equivariance and high-frequency detail fidelity in traditional generative models through deep integration of signal processing theory and deep learning methodologies. The synthesis module consists of three sequential components:

(1) Fourier Feature Input Layer: Introduces Fourier feature mapping as the network's initial input, enhancing translation/rotation equivariance while eliminating grid-aligned artifacts. Its mathematical formulation is as follows:

$$Z_0(x) = \sum_{k=1}^{N_f} \left[a_k \sin(2\pi f_k \cdot x) + b_k \cos(2\pi f_k \cdot x) \right], \tag{2}$$

where x represents the spatial coordinates of each pixel, and parameter f_k is a preset frequency component.

(2) Anti-Aliasing Synthesis Layers: Fourteen cascaded synthesis layers progressively refine feature spaces through rigorous frequency-domain control, enhancing robustness in generation and reconstruction tasks.

(3) ToRGB Output Layer: Decodes high-dimensional features into RGB values that comply with image spatial constraints, directly translating abstract representations to pixels.

Each synthesis layer incorporates a preceding Affine Transformation Layer that converts W+ space encodings to StyleSpace through:

$$s = A(\omega) \odot \gamma + \beta. \tag{3}$$

Here, A represents a fully connected network, while γ and β are learnable gain and bias terms. This operation enables semantic decoupling from the disentangled latent space W to StyleSpace S, permitting independent control of visual attributes (e.g., pose, texture, illumination) across network hierarchies.

4.3 Crypto-Mask

To achieve globally unique style editing, we propose a cryptographically constrained matrix modulation method named **Crypto-Mask**, as illustrated in Figure 3. This approach utilizes hash of user information to directly influence the neural network's internal parameters. By applying this technique to proper weight matrices in the network, we ensure that each user experiences a uniquely personalized visual effect. The technical details are as follows:

Generation of Crypto-Mask. Given the SHA256 hash, denoted as $h_{\text{user}} \in \{0,1\}^{256}$, we expand it into a binary mask matrix M_{mask} tailored to match the dimensions of the target neural network's weight matrix. Assume the weight matrix is $W \in \mathbb{R}^{m \times n}$. Then, the mask matrix is defined as follows:

$$M_{\text{mask}}[i,:] = \begin{cases} 1, & \text{if } h_{\text{user}}^{(i \bmod 256)} = 1 \\ 0, & \text{if } h_{\text{user}}^{(i \bmod 256)} = 0 \end{cases}. \tag{4}$$

This **Bit-Row Exact Mapping** (**BREM**) mechanism ensures that each bit of the hash value directly controls the modulation of the corresponding row of the weight matrix. By using a cyclic mapping, the hash bits are reused when the matrix has more than 256 rows, thereby maintaining encrypted uniqueness across networks of different scales.

Weight Matrix Editing. After generating the binary mask matrix, we apply a scaling operation to edit the original weight matrix. The adjusted weight matrix is given by

$$W_{\text{edited}} = W \odot (\gamma_1 \cdot M_{\text{mask}} + \gamma_0 \cdot (1 - M_{\text{mask}})), \tag{5}$$

where γ_1 and γ_0 are predefined scaling factors applied to the rows where the hash bits are 1 and 0 respectively, and M_{mask} serves as a control matrix to selectively scale specific rows.

This row-based modulation introduces a secure and personalized style alteration within the network's internal representation. Proper selection of γ_1 and γ_0 values, typically in the range $[0.5, 1.5]$, ensures semantic consistency while creating unique visual characteristics.

StyleSpace Perturbation. As demonstrated in the StyleSpace analysis by Wu et al. [91], StyleSpace (S-space) offers greater disentanglement compared to W+ space in the context of StyleGAN. Therefore, it is ideal for applying Crypto-Mask-edited weight matrices. The perturbation in S-space is calculated using the following operations:

$$\Delta s = M \cdot W_{\text{edited}}, \tag{6}$$

$$s_{\text{edit}} = s_{\text{orig}} + \Delta s, \tag{7}$$

where $M \in \mathbb{R}^{d \times n}$ is a learnable mapping matrix, s_{orig} represents the original style vector in S-space, s_{edit} denotes the style vector after applying the perturbation.

This controlled perturbation ensures a balanced tradeoff between preserving the original visual intent and introducing personalized stylistic variations. The cryptographically bound weight adjustments thus act as a secure and verifiable means of NFT customization, fostering on-chain NFT uniqueness.

4.4 NFT Verification

Figure 4 illustrates the NFT asset authenticity verification process within the UniqueNFT framework, which encompasses two modes: regular verification and storage failure recovery. In a blockchain environment, the triggering conditions for NFT asset verification include ownership transfer transactions, user-initiated on-chain arbitration requests, and periodic system checks of storage status. The verification logic of current mainstream NFT architectures heavily relies on the availability of off-chain storage services. The typical process can be summarized as follows: the user wallet address locates the metadata IPFS hash recorded in the smart contract, then the IPFS network is accessed to retrieve the metadata file and parse the resource locator, and finally, a request is sent to the target server to return the actual image file. However, this process suffers from multiple structural flaws.

According to the 2023 industry analysis report by DappRadar, over 60% of NFT projects point image resource locators in their metadata to URL links hosted on centralized cloud services such as AWS and Google Cloud, rather than IPFS hashes based on content addressing. This phenomenon

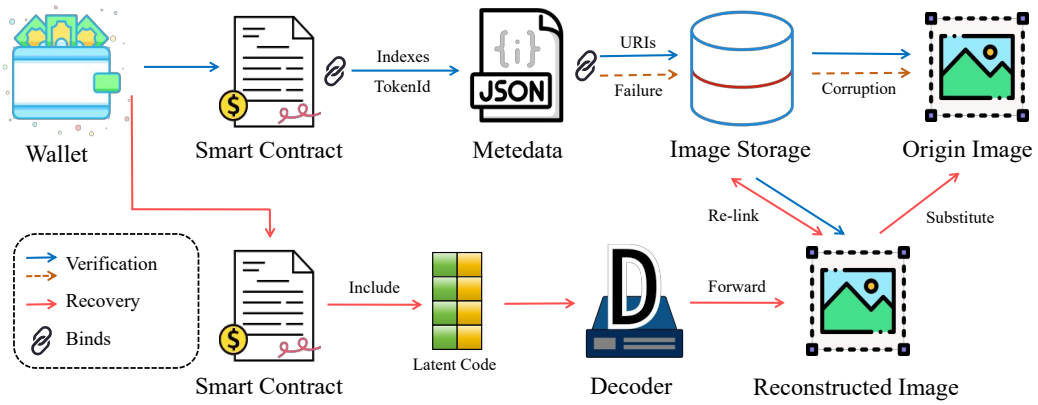


Fig. 4. Process of NFT image integrity verification and semantic recovery with UniqueNFT.

exposes a deviation from the decentralized storage principle in existing solutions. Even when resources are stored on the IPFS network, their persistence still relies on the voluntary pinning behavior of nodes. Any disruption in the storage service, such as IPFS gateway timeouts or centralized server downtime, will directly lead to verification failures, causing a decoupling effect between ownership records and physical assets.

4.4.1 Storage-independent NFT Verification. The UniqueNFT framework combines on-chain semantic encoding with a trusted oracle network to achieve asset verification capabilities that do not rely on traditional storage mediums, such as IPFS, AWS, or other distributed and centralized storage systems. When the system triggers asset verification, the smart contract reads the on-chain stored semantic vectors and hash of user information. These encoded data and model version hashes are then transmitted to an off-chain trusted execution environment via the decentralized blockchain oracle network. Oracle nodes load the corresponding version of the StyleGAN decoder based on the model hash registered in the smart contract and perform the image reconstruction task within the hardware-level isolated trusted execution environment. Subsequently, the oracle node computes the cryptographic hash of the reconstructed image and digitally signs it. The signature result is returned to the on-chain contract for comparison. If the hash matches the asset fingerprint recorded during the initial minting phase, the verification passes; if there is a discrepancy, an exception handling protocol is triggered.

While the UniqueNFT framework provides a storage-independent NFT verification model, integrating IPFS can further reduce gas costs and access latency. IPFS's content addressing and decentralized structure ensure high data integrity, while minimizing the need to re-transmit all data during each verification process.

4.4.2 Storage Failure Handling and Recovery. In scenarios where UniqueNFT utilizes IPFS for storage, UniqueNFT has designed a recovery mechanism to address scenarios where IPFS nodes are lost or external storage becomes unreachable. When decentralized probe nodes detect that IPFS or external storage is unavailable, the system automatically initiates the recovery process. The oracle network uses the on-chain stored semantic vectors to reconstruct high-fidelity images, which are then re-injected into the IPFS network and assigned new content identifiers. Simultaneously, the metadata record in the smart contract is updated. The storage recovery proof generated by this process is permanently recorded on-chain after being compressed using zero-knowledge proof techniques, providing verifiable credentials for subsequent audits.

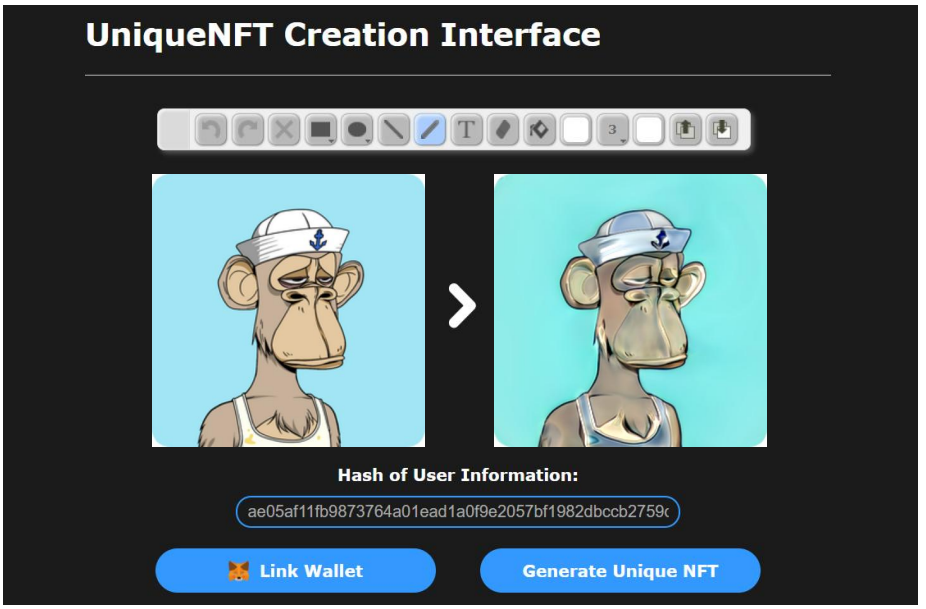


Fig. 5. UniqueNFT creation interface.

Unlike traditional solutions, which passively depend on the availability of storage services, UniqueNFT endows the system with inherent resilience to decay. This technical path provides a theoretical paradigm and practical foundation for the persistent verification of digital assets in the Web3 era.

4.5 User Interaction Interface

In this study, we developed a web-based graphical user interface to support the end-to-end execution of the proposed UniqueNFT framework. As illustrated in Figure 5, the interface referred to as the **UniqueNFT Creation Interface**, enables users to complete the core operations involved in the personalized NFT generation workflow. Within the system, users can upload their original digital image (e.g., artistic illustrations or avatar photos) via the image upload area located on the left side of the interface. The input field labeled **Hash of User Information** is used to enter the hash value corresponding to the user's identity, which serves as the driving variable for the personalization algorithm. By clicking the **Link Wallet** button, users can connect their Ethereum wallet to bind their on-chain identity. Subsequently, clicking the **Generate My NFT** button initiates the personalized generation process, with the resulting image rendered in real time in the right display panel.

In the current version, we employ StyleGAN3 as the backbone image generator, combined with **Encoder for Editing** (**E4E**) as the image inversion module, to construct the complete personalized editing pipeline. All input and output images are standardized to a resolution of 256 × 256 to accommodate both the model architecture and front-end display requirements. As detailed in Section 5.3, we select the final synthesis convolutional layer in StyleGAN3 as the injection point for Crypto-Mask. This design enables fine-grained style modulation while preserving semantic consistency, achieving a balance between generation quality and visual uniqueness.

Overall, the user interaction interface encompasses the full set of operational stages in the UniqueNFT framework, including identity hashing, image uploading, personalized NFT generation,

Table 2. Core Data Fields Specification in UniqueNFT Smart Contract

Field	Type	Usage Phase	Technical Specification
assetHash	bytes32	Minting	Cryptographic hash (SHA3-256) of final NFT media asset
mintBlock	uint256	Auditing	Block number recording mint transaction timestamp
encoderHash	bytes32	Training	Keccak-256 digest of GAN encoder model binary data
decoderHash	bytes32	Training	Keccak-256 digest of styleGAN decoder model parameters
latentCode	uint16[]	Training	Quantized feature vector in Q1.15 fixed-point format
styleCode	bytes	Crypto-Mask	256 bit hash of user information (AES-256-GCM)
cryptoAlgo	string	Security	Cryptographic algorithm identifier with version tag
oracleNode	address	Validation	Certified decentralized oracle node address
taskRegistry	mapping	Management	State tracking mapping for oracle job lifecycle
modelVersion	bytes32	Security	Active model version cryptographic fingerprint
versionLog	mapping	Archival	Immutable registry of historical model hashes

and result visualization. It demonstrates the framework's end-to-end usability, ensures practical ease of use, and serves as a prototype for potential deployment in real-world Web3 applications.

4.6 Contract Design

The UniqueNFT protocol enhances the ERC-721 standard through OpenZeppelin's audited implementation, inheriting its proven NFT ownership management while redefining semantic data preservation.

As detailed in Table 2, the contract introduces eleven state variables that transform NFT metadata management. The assetHash field anchors content authenticity through SHA3-256 hashing, while latentCode employs Q1.15 fixed-point encoding to compress 512D feature vectors - achieving 18 × storage efficiency over conventional FP32 formats without precision loss. Cross-chain interoperability is enabled through oracleNode, which coordinates decentralized validators for media restoration.

The protocol's functional advances are demonstrated through eight core operations (Table 3). The mintSemanticNFT function implements tripartite validation: IPFS storage proofs ensure media persistence, vector dimension checks prevent model mismatches, and model hashes enforce cross-chain version alignment. Governance functions setModelHashes and updateModelVersion establish dual authorization for model updates, with all historical versions immutably logged in versionLog.

Three fundamental innovations distinguish our architecture from conventional NFT standards:

(1) Cryptographic Model Anchoring. Dual hashes (encoderHash/decoderHash) immutably bind ML models to NFTs via **Decentralized Autonomous Organization** (**DAO**) controlled updates, preventing unauthorized model substitution.

(2) Privacy-preserving Parameter Storage. AES-256-GCM encrypted styleCode with dynamic cryptoAlgo identifiers enables user-controlled decryption via decryptStyleCode, eliminating centralized key custodians.

(3) Cross-chain Media Restoration. The requestMediaRecovery function initiates decentralized media regeneration through certified oracles, with recovery states tracked in taskRegistry and verified against original assetHash.

Maintaining full ERC-721 compatibility, the protocol seamlessly integrates with existing NFT marketplaces and wallets while introducing next-generation capabilities for persistent semantic preservation. By combining cryptographic primitives, decentralized verification mechanisms, and

Table 3. Core Function Specifications in UniqueNFT Smart Contract

Function	Parameters	Returns	Technical Specification
setModelHashes	(bytes32 encoderHash, bytes32 decoderHash)	void†	Stores new model hashes in encoderHash and decoderHash fields after DAO multisignature approval
updateModelVersion	(bytes32 newVersionHash)	void†	Activates new model version by updating modelVersion field with governance validation
requestMediaRecovery	(uint256 tokenId)	void†	Initiates recovery workflow through oracleNode address with payment verification
verifyModelHash	(bytes32 targetEncoder, bytes32 targetDecoder)	bool	Validates consistency between input hashes and current encoderHash and decoderHash fields
retrieveLatentVector	(uint256 tokenId)	uint16[]	Returns latentCode field containing 512D Q1.15 fixed-point encoded feature vector
mintSemanticNFT	(address recipient, bytes32 assetHash, uint16[] latentVector, bytes encryptedStyle, bytes ipfsProof)	uint256	Executes tri-phase verification: assetHash proof check, latentVector dimension validation, model version compatibility
getRecoveryStatus	(uint256 tokenId)	enum Status	Queries taskRegistry mapping to return recovery process state (Pending/Completed/Failed)
decryptStyleCode	(uint256 tokenId, bytes privateKey)	int128[]	Decrypts styleCode field using algorithm specified in cryptoAlgo field with user key

†Solidity functions without explicit return values are denoted as void-type.

Fig. 6. Bored ape yacht club datasets.

highly efficient model compression, the design embodies a cohesive and forward-looking engineering philosophy. Through this reimagining of on-chain metadata logic, UniqueNFT establishes a robust and extensible foundation for secure, efficient, and future-proof digital asset management.

5 Experiment

5.1 Dataset

To evaluate the effectiveness of our personalized editing approach, we selected the **Bored Ape Yacht Club (BAYC)**[5] dataset on HuggingFace,[6] one of the most iconic NFT datasets (as shown in Figure 6). Comprising 9,999 images, BAYC is characterized by its rich textures and high color complexity. Compared to other NFT datasets, its visual diversity and intricate details make it an ideal choice for assessing the performance of personalized editing methods.

[5]https://boredapeyachtclub.com
[6]https://huggingface.co/huggingnft

5.2 Model Training

Training Hardware. The experiments were conducted on a hardware platform equipped with an NVIDIA GeForce RTX 3090 GPU (24 GB GDDR6X VRAM), an Intel Xeon Platinum 8362 processor, and 45 GB DDR4 RAM. The Ubuntu 18.04 LTS (Bionic Beaver) operating system was employed to ensure compatibility with CUDA acceleration libraries and deep learning frameworks.

Encoder Training Configuration. The E4E encoder [3, 4], built on a ResNet progressive backbone network [41], processed 6-channel preprocessed data. It was trained for 500,000 iterations using the Ranger optimizer (a hybrid of RAdam and Lookahead strategies) with an initial learning rate of 0.0001 and a batch size of 4. The composite loss function included an L2 reconstruction loss (λ_1=1.0) for pixel-level accuracy, an LPIPS [97] perceptual loss (λ_2=0.8) to preserve high-level semantic similarity, and a MoCo contrastive loss [40] (λ_3=0.5) to enhance latent space coherence. The encoder-generated latent vectors were decoded into 256×256 resolution images via a frozen pre-trained StyleGAN3-R generator, initialized with truncated W-space mean values (Ψ=0.7). Notably, identity preservation loss was disabled due to the homogeneous character features in the Bored Ape NFT dataset, eliminating redundant constraints on stylized yet structurally consistent artwork.

Decoder Training Configuration. The StyleGAN3 generator [45] adopted the R architecture and was trained using the Adam optimizer [49] (momentum coefficients β_1=0, β_2=0.99; numerical stability constant ϵ =1e-8). The generator and discriminator learning rates were set to 0.0025 and 0.0015, respectively, with a batch size of 8 to accommodate GPU memory constraints. Training spanned 5,000 **kilokernel images** (**kimg**) to ensure comprehensive distribution coverage. To mitigate mode collapse, we incorporated an R1 gradient penalty [56] (regularization weight λ=6) to constrain the L2 norm [77] of the discriminator's gradients on real samples, alongside a logistic loss function and **adaptive discriminator augmentation** (**ADA**) to dynamically stabilize training.

5.3 Optimal Layer Perturbation for NFT Stylization

This section systematically evaluates the layer sensitivity of the Crypto-Mask perturbation mechanism applied to the affine weights within the synthesis layers of the generative network for NFT image editing. The experiment utilizes ten standard images from the Bored Ape Yacht Club collection as baseline samples and incorporates six distinct user information hashes derived from a blockchain network as deterministic editing parameters.[7] A layer-wise perturbation strategy is employed, wherein Crypto-Mask modifications are applied separately to each of the 15 synthesis layers (Layer 0–14) of StyleGAN3 during the generation process. In total, 900 edited samples are generated to comprehensively examine the effects of perturbations across different layers. For the Crypto-Mask perturbation, the parameters are set with $\gamma_1 = 1.4$ and $\gamma_0 = 0.6$ to control the strength and range of the perturbations applied to the weight matrices.

To quantitatively assess the impact of Crypto-Mask across layers, a dual-metric evaluation framework is established. First, for a given image edited at a specific layer, the **structural similarity index** (**SSIM**) is computed across all pairwise comparisons among the six user-generated results, with the average SSIM serving as a measure of inter-user editing divergence. The SSIM value ranges from [0,1], where 0 indicates complete dissimilarity and 1 signifies identical outputs. Lower SSIM values suggest that the given layer enables more pronounced stylistic diversity. Second, the **peak signal-to-noise ratio** (**PSNR**) is calculated between the edited images and their original unaltered versions, with the average PSNR at each layer representing the semantic fidelity of the

[7]The six user information hashes are: a3fa708e129065c4ae4de0b796d1593969a3af7e512913ec7124cef94f7cb11b, 1b483fbf4eb272bcef84eb14d9719f38b01a5f2737b231ad659c10507570b066, 170623dcf3957b6a6d9824280c863cb0a7e6e9ad6642c49e98e4e9634d0f1fc4, 4ebbc4ee0833ca47c516dcf24fa916389f1b4b4a8cceeb9a130011a6f8fd8d76, ae05af11fb9873764a01ead1a0f9e2057bf1982dbccb2759dbb6a970dcd36c2b, 51ee43485bc59ec0bae5bdbea4a25fb35fd43ea11b0ab90148ef61ecf62f6afd.

Table 4. Average SSIM and PSNR Values Across Different Synthesis Layers

Layer	Image Sample										Avg SSIM	Avg PSNR
	1	2	3	4	5	6	7	8	9	10		
Layer 0	0.9594	0.9332	0.9162	0.9583	0.9004	0.9293	0.9296	0.9126	0.9406	0.9422	0.9322	20.36
Layer 1	0.9455	0.9234	0.9021	0.9477	0.8911	0.9190	0.9200	0.8958	0.9229	0.9213	0.9189	19.94
Layer 2	0.9420	0.9269	0.9020	0.9437	0.8770	0.9186	0.9214	0.8978	0.9169	0.9201	0.9166	20.07
Layer 3	0.9161	0.8798	0.8635	0.9191	0.8363	0.8981	0.8800	0.8624	0.8901	0.8856	0.8831	20.06
Layer 4	0.9382	0.9136	0.8934	0.9374	0.8793	0.9036	0.9138	0.8793	0.9105	0.9101	0.9079	20.28
Layer 5	0.9099	0.8799	0.8677	0.9175	0.8337	0.8804	0.8772	0.8440	0.8834	0.8866	0.8781	20.42
Layer 6	0.8723	0.8336	0.8126	0.8831	0.7898	0.8228	0.8275	0.7901	0.8419	0.8454	0.8319	19.69
Layer 7	0.8886	0.8503	0.8363	0.8988	0.8019	0.8333	0.8435	0.7921	0.8584	0.8638	0.8467	20.06
Layer 8	0.8562	0.8323	0.8067	0.8684	0.7788	0.7971	0.8224	0.7658	0.8282	0.8240	0.8180	19.56
Layer 9	0.8709	0.8435	0.8214	0.8878	0.7938	0.8175	0.8362	0.7791	0.8369	0.8495	0.8337	19.93
Layer 10	0.8354	0.8031	0.7801	0.8585	0.7716	0.7681	0.7838	0.7265	0.8029	0.8080	**0.7938**	**19.22**
Layer 11	0.8703	0.8550	0.8417	0.8922	0.8120	0.8397	0.8371	0.7887	0.8452	0.8578	0.8440	19.99
Layer 12	0.7607	0.7315	0.7529	0.7599	0.7190	0.7239	0.7354	0.6626	0.7400	0.7192	**0.7305**	**18.75**
Layer 13	0.8397	0.8628	0.8563	0.8485	0.7972	0.8538	0.8441	0.7760	0.8559	0.8121	0.8346	19.42
Layer 14	0.8048	0.7968	0.8210	0.8420	0.7822	0.7464	0.7949	0.7140	0.7656	0.8260	**0.7894**	**17.52**

perturbations. The PSNR metric ranges from $[0, \infty]$, where higher values indicate less structural degradation due to editing operations. These complementary metrics enable a precise evaluation of the tradeoff between stylistic diversity and semantic preservation across different layers.

According to the data in Table 4, the average SSIM values for the first five layers are generally above 0.88, while PSNR values remain stable within the 20.0–21.4 range. This indicates that applying the Crypto-Mask at these layers fails to induce significant visual feature variations, rendering its editing effect below the perceptual threshold (JND, Just Noticeable Difference). A sharp decline in SSIM (0.8319) and PSNR (19.69) is observed starting from Layer 6, marking the activation of an effective editing window. Despite minor fluctuations, the overall trend exhibits a positive correlation between layer depth and metric degradation. To establish an objective selection criterion, this study sets an average SSIM threshold of 0.8 to define significant editing responses, as layers with lower SSIM values are more likely to be key controllers of image style. Based on this criterion, we identify Layer 10 (0.7938), Layer 12 (0.7305), and Layer 14 (0.7894) as primary targets for further analysis. Notably, although all three layers meet the SSIM threshold, their editing characteristics differ significantly: Layer 10 exhibits a perceptual PSNR gap of 1.7 dB compared to Layer 14 (19.22 vs. 17.52), whereas Layer 12 experiences a drastic 13.5% drop in SSIM (0.7305) relative to adjacent layers, forming a statistically significant outlier.

Visualization analysis (shown in Figure 7) reveals that while Layer 10 perturbations meet statistical significance (SSIM threshold), they only induce minor localized texture jittering, failing to achieve an obvious visual style transformation. In addition, severe pixel-level disorder reorganization occurs in all the images generated by the Layer 12 perturbation, making the image fidelity distorted. This may be due to the incomplete decoupling of features controlled by Layer 12, resulting in a collapse of structural consistency when perturbation is performed at this layer, making the generated images unavailable.

Summary. In comparison, Layer 14 demonstrates the most optimal balance of editing characteristics. As indicated in Table 4, this layer maintains semantic recognizability (PSNR = 17.52) while

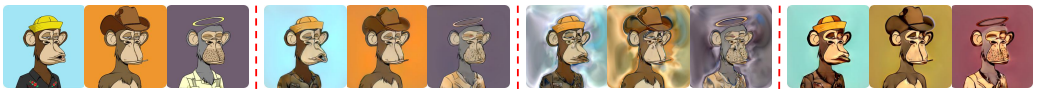


Fig. 7. **Multilayer editing visualization contrast.** Leftmost: Source NFT specimens. Adjacent columns show sequential edits in Layers 10 (mild feature shifts), 12 (semantic fractures in facial geometry), and 14 (coherent style transfer with anatomical preservation).

achieving substantial inter-user diversity (SSIM = 0.7894). As illustrated in Figure 7, its perturbations primarily affect stylistic attributes such as hair texture glossiness and background detail complexity, without compromising the subject's identity. This superior style modification capability enables the Crypto-Mask to generate aesthetically diverse NFT variants while preserving their core value, striking a perfect balance between creative editing and structural fidelity.

5.4 Crypto-Mask Effect by Multi-User

This section presents case studies of Crypto-Mask applied to seven original NFT images (Figure 8), employing the six user information hash values utilized in the Section 5.3 for consistency, to provide an intuitive verification of Crypto-Mask's effectiveness in controlled generation and unique stylistic personalization. By encoding these hashes as perturbation directives for the Affine transformation weights in the 14th layer of StyleGAN3, Crypto-Mask ($\gamma 1 = 1.4, \gamma 2 = 0.6$) successfully demonstrates a strong binding effect between user identity and generative style. The results highlight both the consistency of generated styles within the same user and the personalization of NFTs across different users.

Intra-User Style Consistency. Experimental results reveal that images generated using the same user information hash exhibit highly consistent stylistic characteristics across different NFT prototypes. Specifically:

- **User1** produces NFTs with smooth edge transitions (e.g., refined hat and fur boundaries), vibrant warm-toned backgrounds, and consistently closed eyes across all samples.
- **User2** retains the original color scheme while introducing a distinctive swollen deformation around the left eye, alongside a granular texture on clothing.
- **User3** generates NFTs with a dark-toned aesthetic, incorporating grayscale patches on the eyelids and lips with reduced surface reflections.
- **User4** introduces high-saturation orange regions on both the hat brim and lips while employing a bright monochromatic background.
- **User5** adopts a cool-toned gray-blue palette, producing a uniform color distribution with an ink-wash-like diffusion effect.
- **User6** exhibits a dramatic stylization, adding a white highlight in the left eye region and significantly increasing the curvature of mouth lines.

These findings indicate that user's information hashes are effectively encoded as stable style control directives through the Crypto-Mask mechanism. This allows a user's personalized visual signature to remain stable across various prototype modifications. Unlike traditional NFT minting, which relies solely on hash values for uniqueness, Crypto-Mask transforms each user information hash into a distinctive visual grammar. This ensures differentiation between users' NFTs from the same base image while allowing consistent style inheritance across multiple works. Such properties align perfectly with web3 applications, where users seek both individuality and coherent digital identity representation. Consequently, Crypto-Mask opens a new dimension for personalized expression in blockchain-based digital assets.

Fig. 8. **Crypto-Mask editing effects.** For a set of 7 input samples, user-specific mask perturbations derived from 6 distinct blockchain identity hashes are applied to the affine transformation module of the 14th synthesis layer during the image decoding phase, demonstrating consistent stylization within each user group and significant divergence across users.

Inter-User Visual Differentiation. By analyzing Figure 6, which compares different users' edited results from the same base image, we observe five key dimensions of inter-user differentiation:

- **Facial and Fur Tones**: In the first and second rows, User1, User2, and User6 maintain a lifelike flesh-tone, whereas User3 and User5 shift toward a grayish-white palette.
- **Eye State**: In the fourth row, originally open-eyed samples close their eyelids after editing by User1, User2, User3, and User4, while User5 and User6 retain the open-eye state. Notably, User2 exhibits a distinctive swollen deformation in the left eye region.
- **Mouth Features**: In the third row, where the original ape holds a cigarette, User2, User4, and User6 completely remove the cigarette, whereas others retain it with color adjustments.

In the fifth row, which features a puckered-lip expression, User3 eliminates the expression, while other users preserve it.

— **Hat and Accessories**: In the second row, the sailor hat loses its original logo and takes on a multi-colored appearance for User1 and User3, while Users 2, 4, 5, and 6 retain the hat shape but shift its color to white-gray or white-blue.

— **Background**: All user groups deviate from the original monochrome background, producing diverse high-contrast transitions. User1's samples feature a glowing central area, while User5's incorporate layered blue-gray tones.

In conclusion, our experimental results demonstrate that the Crypto-Mask mechanism effectively maps user identity hashes to distinct and controllable style transformations. Different hashes of user information influence not only localized attributes (e.g., eye state, mouth shape, and clothing color) but also the overall visual style (e.g., color temperature, lighting effects, and texture properties). This mechanism ensures that the uniqueness of an NFT is no longer solely dependent on the hash assigned at the time of minting but is further refined by incorporating hash of user information to precisely control the generated style. As a result, each NFT exhibits a distinct visual identity while maintaining stylistic consistency across a given user's NFT collection. This novel approach introduces a new identity-mapping paradigm in the NFT space, potentially transforming NFTs from static collectibles into highly customizable and expressive digital assets, thereby enhancing their applicability in the Web3 and blockchain ecosystems.

6 Evaluation

6.1 Performance Comparison of Personalized Editing Techniques

In this section, we evaluate the performance of two Crypto-based personalized image editing methods–Crypto-Mask and Crypto-Dropout [30, 31], by comparing their impact on image generation quality.

Experimental Design. We selected 15 images from the Bored Ape Yacht Club dataset as evaluation samples and chose 6 different users. Each hash of user information was used to control the network parameters during the image generation process. Specifically, Crypto-Dropout leverages the hash value to control the "dropout" operation in the 14th synthesis layer's convolutional layers (where a hash value of 0 indicates dropout and 1 indicates retention). In contrast, Crypto-Mask uses the hash value to modulate the gain of the affine weight matrix in the same layer. For each image, we applied the hash values of 6 users to control the image generation under both methods and compared the results with the original images.

To quantitatively assess the effectiveness of both methods, we computed the PSNR and SSIM values between each generated image and the corresponding original image. PSNR [37], a traditional image quality metric, measures the similarity in signal space; higher PSNR values indicate less signal deviation between the generated and original images. SSIM [86], on the other hand, focuses on structural similarity, offering a closer approximation to human perception of image quality. By comparing these two metrics, we were able to comprehensively evaluate the performance of each method in preserving image details and structural integrity.

Quantitative Analysis. The experimental results are summarized in Figure 9. From the global statistics, it is evident that Crypto-Dropout significantly underperforms Crypto-Mask, with an average PSNR of 11.85 and SSIM of 0.31 for Crypto-Dropout, compared to 17.28 for PSNR and 0.78 for SSIM in the case of Crypto-Mask. Specifically, the generated images with Crypto-Dropout consistently show lower values for both metrics. According to the calculation formulas for PSNR and SSIM, the lower PSNR value suggests reduced signal similarity, while the lower SSIM indicates substantial structural loss. These results imply that the Crypto-Dropout method causes considerable

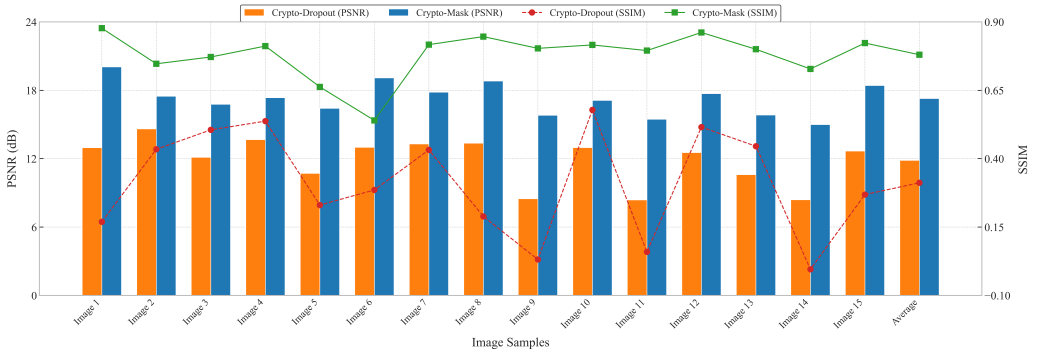


Fig. 9. Quality comparison of customized images: Crypto-Mask vs. Crypto-Dropout.

Fig. 10. Visual comparison of personalized image editing: Crypto-mask vs. Crypto-dropout. This figure presents the results of applying two different personalization algorithms to three image samples using hash information from four of the six experimental users.

damage to the semantic and structural content of the images, leading to a noticeable decline in image quality, with significant blurring and distortion.

Visual Analysis. As shown in Figure 10, the images generated with Crypto-Dropout exhibit significant detail loss and distortion. In particular, the contours of the central character's face and clothing textures appear rough, and certain areas suffer from visible pixelation effects–where large sections of the image lose detail, creating coarse block-like regions. This results in an overall poor visual impression. Key semantic information, such as the character's facial features and object outlines, becomes blurred or lost, severely affecting the quality and usability of the generated image. These visual discrepancies align with the low PSNR and SSIM values observed in the data analysis, corroborating the hypothesis that Crypto-Dropout's coarse manipulation leads to substantial degradation of the image's integrity. In contrast, images generated using Crypto-Mask maintain much more of the original structure and detail. Even under personalized modification, key elements such as facial features and clothing textures remain clear and intact, with the visual quality of the generated images significantly surpassing that of Crypto-Dropout.

Summary. Based on both the quantitative evaluation and visual analysis, it is clear that the Crypto-Mask method significantly outperforms Crypto-Dropout in terms of image quality and fidelity. While both methods personalize images using user hash values, the Crypto-Dropout method's aggressive approach results in substantial loss of detail, structure, and semantic

information, leading to a marked decline in the image's overall quality and usability. On the other hand, Crypto-Mask leverages fine-tuned affine weight adjustments to retain crucial image details and structural elements during personalization, generating higher-quality images that are more usable and visually compelling. We believe that the Crypto-Mask method demonstrates a clear advantage for personalized NFT generation tasks and holds promise for applications in blockchain-based digital asset scarcity protection and personalized customization, driving innovation in the field of personalized digital art.

6.2 User Study

6.2.1 Experiment Design. This experiment is to evaluate user preferences for two personalized image editing algorithms, Crypto-Mask and Crypto-Dropout, as well as their satisfaction with images generated by the Crypto-Mask algorithm.

Participant Information. A total of 55 participants were involved in this experiment, with 63.64% male and 36.36% female. The age distribution is as follows: 4 participants (7.27%) are under 18 years old, 40 participants (72.73%) are between 18 and 24 years old, 8 participants (14.55%) are between 25 and 34 years old, and 3 participants (5.45%) are over 34 years old. All participants had no known visual impairments and voluntarily participated in the experiment. Participant data was anonymized to ensure privacy throughout the process.

Experimental Setup. The experiment was conducted using a questionnaire format, with the following main sections:

(1) Visual Detectability and Difference Recognition in Crypto-Mask Edits. Participants were asked to assess whether they could distinguish between the Crypto-Mask edited images and the original images and to identify specific visual differences (e.g., background color, accessories on the Ape, features of the Ape itself such as fur, color, and expression).
(2) Image Preference Comparison. Participants were asked to compare images generated by Crypto-Mask and Crypto-Dropout and select their preferred images.
(3) Overall Quality Satisfaction Rating for Crypto-Mask Images. Participants rated their satisfaction with the images on a scale of 1 to 5, where 1 indicated "very dissatisfied" and 5 indicated "very satisfied."

Data were collected through participants' ratings of each image, generating quantitative data on preferences, difference recognition, and satisfaction.

6.2.2 Results and Analysis. Figures 11, 12, and 13 show user preferences for different algorithms, satisfaction ratings for Crypto-Mask edited images, and results of the visual difference perception survey.

Visual Distinction Ability (Figure 11). According to the survey results, 98.18% of participants (54 out of 55) were able to identify differences between the Crypto-Mask-edited images and the original versions, while only 1.82% (1/55) failed to detect any change. A one-sample chi-square test confirmed that this identification rate was significantly higher than the 50% baseline expected by random chance ($\chi^2 = 52.36$, df = 1, $p < 0.0001$). Among the three edited feature groups, users demonstrated the highest recognition rate for changes to the "Ape's core features" (95.2%), significantly outperforming recognition rates for background color changes (81.8%) and accessory modifications (76.4%). The inter-group differences were statistically significant ($\chi^2 = 12.54$, $p = 0.002$), indicating that Crypto-Mask achieves perceptually salient edits beyond low-level pixel noise.

Image Preference Comparison (Figure 12). As shown in Figure 12, participants consistently preferred images generated with Crypto-Mask over those produced by Crypto-Dropout. The overall selection ratio was 258:72 (78.2% vs. 21.8%), with a highly significant deviation from a uniform distribution ($\chi^2 = 128.89$, df = 1, $p < 0.0001$). The preference rates in individual pairs ranged from

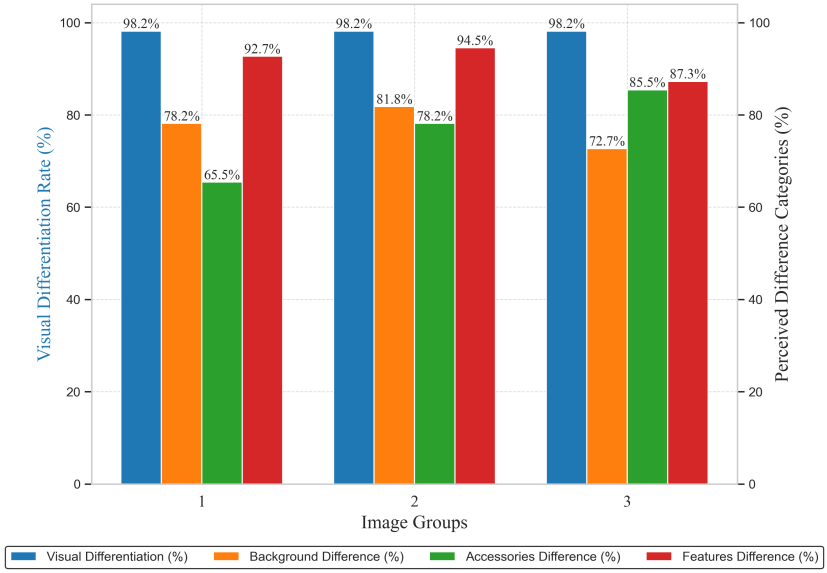


Fig. 11. Perception accuracy and difference recognition by image group.

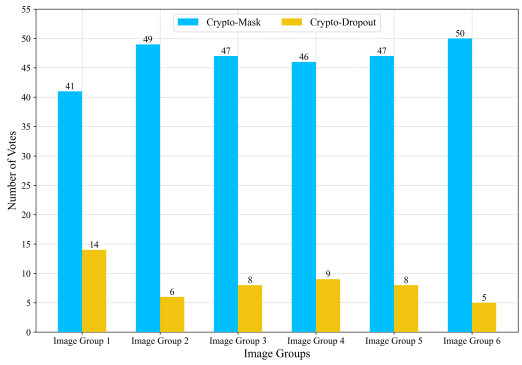


Fig. 12. Comparison between crypto-mask and crypto-dropout.

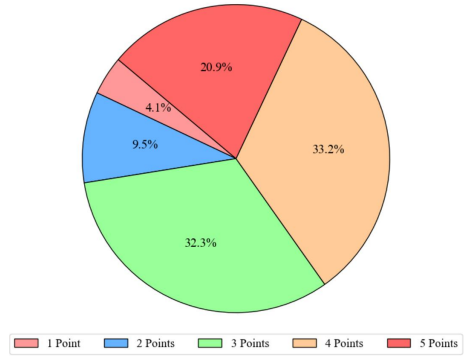


Fig. 13. Satisfaction distribution for crypto-mask images.

74.5% to as high as 90.9%, with all groups showing statistically significant preferences ($p < 0.001$, chi-square test). This consistent pattern suggests a systematic user preference for Crypto-Mask-generated images rather than random variation.

Satisfaction Rating Distribution (Figure 13). Satisfaction ratings for Crypto-Mask-edited images exhibited a significant positive skew. Among 220 valid responses, the mean rating was 3.57 (SD = 0.92), which was significantly higher than the neutral midpoint of 3 (one-sample t-test: $t = 9.19$, $df = 219$, $p < 0.0001$). Specifically, 33.2% of participants rated the images 4 out of 5, and 20.9% gave the highest rating of 5, resulting in a combined high-score proportion (4–5) of 54.1%. A one-sample chi-square test indicated that this proportion did not significantly differ from the expected 50% baseline ($\chi^2 = 1.48$, $df = 1$, $p = 0.224$). Only 13.6% (30/220) of responses fell below 3, suggesting a generally favorable user perception of the image quality.

6.2.3 Summary. The results confirm that the Crypto-Mask algorithm introduces visually perceptible and meaningful edits, with a 98.18% recognition rate among users. It outperforms the baseline Crypto-Dropout method in both user preference (78.2% vs. 21.8%; $\chi^2 = 128.89$, $p < 0.0001$) and satisfaction (mean score = 3.57; $t = 9.19$, $p < 0.0001$). These findings demonstrate Crypto-Mask's effectiveness in addressing user demands for personalized and perceivable image transformations, and underscore its potential utility in real-world image editing applications.

7 Limitations and Future Work

The UniqueNFT framework leverages an innovative protocol design to ensure that NFTs are no longer dependent on external server stability. Instead, it achieves permanence by encoding the semantic representation of NFTs directly onto the blockchain. Additionally, the Crypto-Mask mechanism utilizes hash function properties to guarantee that each user affects the neural network's core layers in a unique manner during image editing, ensuring distinct visual outputs.

Effective On-Chain Verification Mechanism. In the current UniqueNFT framework, all uploaded images undergo Crypto-Mask transformation based on the user's hashed information, ensuring uniqueness. However, for non-artistic, commemorative, or everyday-use images, users may prefer to upload the original picture without modification. To accommodate broader user needs, the framework must support both unaltered and personalized NFT uploads. This, however, introduces a critical challenge: if a personalized NFT is copied and re-uploaded in its original form by an unauthorized party, it could undermine the uniqueness enforced by Crypto-Mask, thereby compromising the protective mechanism. Thus, an effective on-chain verification system is essential to prevent unauthorized image uploads that bypass personalization safeguards. A key consideration is defining rejection criteria based on similarity thresholds. Different categories–such as high-value artistic NFTs, collectible digital art, and casual images–may require distinct similarity evaluation standards with adaptive thresholding. Determining these thresholds and refining the detection framework through extensive empirical research will be a crucial direction for future work.

Enhancing Visual Differentiation. In the UniqueNFT framework, the uniqueness of NFTs is entirely dependent on the Crypto-Mask mechanism, which exhibits strong potential in generating highly diverse and visually distinctive NFTs (as illustrated in Figure 14, where we showcase some of the awesome results from the Random-Mask experiment as a visual demonstration of Crypto-Mask's limitless potential in generating diverse styles). However, in extreme cases, user-specific hash values may exhibit high binary similarity, differing only by a few bits, which could result in minimal perceptible differences in the generated images. To address this, future research can explore more advanced Crypto-Mask designs that ensure significant stylistic variations even when hash values share similar binary distributions, thereby enhancing the visual uniqueness of NFTs. This improvement would not only increase the distinguishability of personalized NFTs but also reinforce the practical value of the UniqueNFT framework in the Web3 era, fostering further advancements in decentralized digital asset innovation.

Strengthening Style Disentanglement. The effectiveness of the Crypto-Mask mechanism hinges on the ability of generative models to disentangle style attributes from semantic content, ensuring that perturbations selectively modify stylistic elements without disrupting core semantics. Ideally, a well-structured generative model would feature distinct layers or modules responsible for controlling tone, texture, and background, allowing Crypto-Mask to apply precise, layer-specific perturbations for controlled and diverse personalization. However, existing generative models still exhibit limitations in style disentanglement. For example, while models like StyleGAN demonstrate certain degrees of feature separation, the extracted attributes often do not align with intuitive human interpretations, restricting their usability for controlled editing. Future advancements should

Fig. 14. **Perturbation effects at the 14th synthesis layer.** For an arbitrary input sample, randomly generated mask perturbations are applied to the affine transformation module of the 14th synthesis layer during the image decoding phase, showcasing the powerful potential of our method in achieving style diversity through layer-specific editing.

focus on improving generative models' ability to recognize and manipulate human-perceivable style features, thereby enhancing the controllability of style modifications. Additionally, exploring independent, model-agnostic style disentanglement modules that can be integrated into various generative architectures to enhance style controllability presents another promising research avenue.

Extension to Multi-type Digital Assets. At this stage, the UniqueNFT framework and its core personalization algorithm have demonstrated strong cross-domain adaptability and transferability within the realm of 2D image-based digital assets. Based on our preliminary attempts on the FFHQ dataset, which differs significantly from NFT images in both style and semantic structure, Crypto-Mask is capable of producing personalized modifications in aspects such as color tone, texture, and facial expression, while preserving the original semantic structure of the images (see Figure 15). Moreover, as shown in Figure 16, Crypto-Mask consistently outperforms Crypto-Dropout in PSNR and SSIM evaluations, both at the single-image level and in overall average metrics, further validating its scalability, generalizability, and superior performance in image-based digital asset scenarios.

Despite its initial success in 2D image domains, the current design of the framework does not yet support more complex types of digital assets, such as video NFTs, 3D models, or multimodal content. These asset types typically involve higher-dimensional temporal structures, spatial representations, and semantic alignments, making them incompatible with the current latent-space mapping and encrypted editing mechanisms. Recent progress in multimodal content generation, fusion and alignment technology [8, 32, 48, 61] provides valuable insights into extending such frameworks beyond visual modalities. Future work may explore several promising directions, including the integration of cross-modal generative models, voxel- or mesh-based 3D style modulation methods, and dynamic masking strategies tailored for sequential data. By incorporating these technologies, the UniqueNFT framework can be further extended to support a broader spectrum of Web3 digital asset scenarios, driving the continued evolution of personalized generation techniques in decentralized environments.

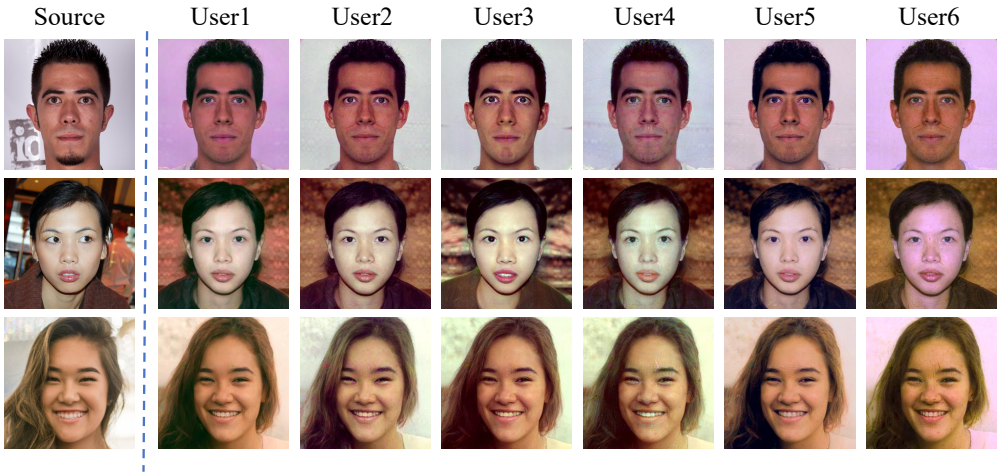


Fig. 15. Crypto-mask editing effects on FFHQ.

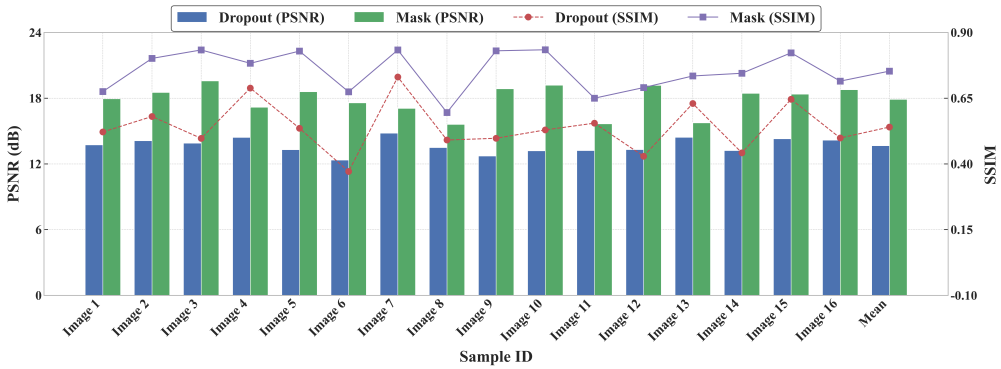


Fig. 16. Quality comparison of customized FFHQ images: Crypto-Mask vs. Crypto-Dropout.

8 Conclusion

This article presents UniqueNFT, a decentralized framework designed to address the challenges of NFT storage persistence and visual uniqueness. By innovatively introducing a decentralized middleware service capable of invoking pretrained encoders and decoders, our framework ensures that NFTs are permanently and securely stored on the blockchain in the form of semantic encodings, allowing for lossless recovery at any time. This fundamentally eliminates reliance on external storage providers and enables NFTs to achieve true decentralized permanence. UniqueNFT is designed with high interoperability and extensibility, allowing it to be seamlessly integrated into existing blockchain and NFT ecosystems. To evaluate the effectiveness of Crypto-Mask mechanismt, we conducted extensive experiments on benchmark NFT datasets, applying this approach and validating its efficacy through quantitative metrics, visualization analyses, and user studies. The results demonstrate its robust effectiveness, usability, and high user satisfaction in generating uniquely stylized NFTs. We firmly believe that UniqueNFT and its core editing mechanism, Crypto-Mask, represent not only a breakthrough innovation in decentralized digital asset protection but also a significant milestone in the evolution of personalized NFTs. By addressing the long-standing issues of NFT storage durability and individual uniqueness, our framework enhances the cultural,

artistic, and economic value of blockchain-based assets. Moving forward, we envision UniqueNFT as a transformative invention in decentralized ecosystems, fostering the next generation of NFT-based Web3 applications, empowering global blockchain communities, and paving the way for a new era of digital art, on-chain identity, and decentralized asset management.

References

[1] Rameen Abdal, Yipeng Qin, and Peter Wonka. 2019. Image2stylegan: How to embed images into the stylegan latent space? In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 4432–4441.

[2] Rameen Abdal, Yipeng Qin, and Peter Wonka. 2020. Image2stylegan++: How to edit the embedded images?. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 8296–8305.

[3] Yuval Alaluf, Or Patashnik, and Daniel Cohen-Or. 2021. Restyle: A residual-based stylegan encoder via iterative refinement. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 6711–6720.

[4] Yuval Alaluf, Or Patashnik, Zongze Wu, Asif Zamir, Eli Shechtman, Dani Lischinski, and Daniel Cohen-Or. 2022. Third time's the charm? image and video editing with stylegan3. In *Proceedings of the European Conference on Computer Vision*. Springer, 204–220.

[5] Omar Ali, Mujtaba Momin, Anup Shrestha, Ronnie Das, Fadia Alhajj, and Yogesh K. Dwivedi. 2023. A review of the key challenges of non-fungible tokens. *Technological Forecasting and Social Change* 187 (2023), 122248.

[6] Rabih Amhaz, Cedric Bobenrieth, and Marlene Marz. 2024. Impact of decentralized autonomous organizations (DAO) on society 5.0. *Proceedings of the Machine Learning, IOT and Blockchain, Dubai, United Arab Emirates* (2024), 17–18.

[7] Andreas M. Antonopoulos. 2014. *Mastering Bitcoin: Unlocking Digital Cryptocurrencies*. " O'Reilly Media, Inc."

[8] Muhammad Ayaz, Mustaqeem Khan, Muhammad Saqib, Adel Khelifi, Muhammad Sajjad, and Abdulmotaleb Elsaddik. 2024. Medvlm: Medical vision-language model for consumer devices. *IEEE Consumer Electronics Magazine* 14, 5 (2024), 75–83.

[9] Seyed Mojtaba Hosseini Bamakan, Nasim Nezhadsistani, Omid Bodaghi, and Qiang Qu. 2022. Patents and intellectual property assets as non-fungible tokens; key technologies and challenges. *Scientific Reports* 12, 1 (2022), 2178.

[10] David Bau, Hendrik Strobelt, William Peebles, Jonas Wulff, Bolei Zhou, Jun-Yan Zhu, and Antonio Torralba. 2019. Semantic photo manipulation with a generative image prior. *ACM Transactions on Graphics (TOG)* 38, 4 (2019), 1–11.

[11] David Bau, Jun-Yan Zhu, Hendrik Strobelt, Bolei Zhou, Joshua B. Tenenbaum, William T. Freeman, and Antonio Torralba. 2018. Gan dissection: Visualizing and understanding generative adversarial networks. arXiv:1811.10597. Retrieved from https://arxiv.org/abs/1811.10597

[12] Juan Benet. 2014. Ipfs-content addressed, versioned, p2p file system. arXiv:1407.3561. Retrieved from https://arxiv.org/abs/1407.3561

[13] Nazanin Zahed Benisi, Mehdi Aminian, and Bahman Javadi. 2020. Blockchain-based decentralized storage networks: A survey. *Journal of Network and Computer Applications* 162 (2020), 102656.

[14] Vitalik Buterin. 2014. A next-generation smart contract and decentralized application platform. *White Paper* 3, 37 (2014), 2–1.

[15] Mehmet Utku Celik, Gaurav Sharma, A. Murat Tekalp, and Eli Saber. 2005. Lossless generalized-LSB data embedding. *IEEE Transactions on Image Processing* 14, 2 (2005), 253–266.

[16] Chef Sperxos. 2021. *Glow in the Dark NFT Official Site*. Retrieved December 14, 2025 from https://opensea.io/collection/glow-in-th-dark

[17] Hongzhou Chen, Haihan Duan, Maha Abdallah, Yufeng Zhu, Yonggang Wen, Abdulmotaleb El Saddik, and Wei Cai. 2023. Web3 metaverse: State-of-the-art and vision. *ACM Transactions on Multimedia Computing, Communications and Applications* 20, 4 (2023), 1–42.

[18] Lei Chen and Shihong Wang. 2017. A secure blind watermarking scheme based on DCT domain of the scrambled image. arXiv:1708.09535. Retrieved from https://arxiv.org/abs/1708.09535

[19] Konstantinos Christidis and Michael Devetsikiotis. 2016. Blockchains and smart contracts for the internet of things. *IEEE Access* 4 (2016), 2292–2303.

[20] Ingemar Cox, Matthew Miller, Jeffrey Bloom, Jessica Fridrich, and Ton Kalker. 2007. *Digital Watermarking and Steganography*. Morgan kaufmann.

[21] Antonia Creswell and Anil Anthony Bharath. 2018. Inverting the generator of a generative adversarial network. *IEEE Transactions on Neural Networks and Learning Systems* 30, 7 (2018), 1967–1974.

[22] Michael Crosby, Pradan Pattanayak, Sanjeev Verma, Vignesh Kalyanaraman. 2016. Blockchain technology: Beyond bitcoin. *Applied Innovation* 2, 6–10 (2016), 71.

[23] Erik Daniel and Florian Tschorsch. 2022. IPFS and friends: A qualitative comparison of next generation peer-to-peer data networks. *IEEE Communications Surveys and Tutorials* 24, 1 (2022), 31–52.

[24] Dipanjan Das, Priyanka Bose, Nicola Ruaro, Christopher Kruegel, and Giovanni Vigna. 2022. Understanding security issues in the NFT ecosystem. ArXiv. arXiv:2111.08893. Retrieved from https://arxiv.org/abs/2111.08893
[25] Maryam Dashti, Reza Safabakhsh, Mohammadreza Pourfard, and Mohammadjavad Abdollahifard. 2015. Video logo removal using iterative subsequent matching. In *Proceedings of the 2015 The International Symposium on Artificial Intelligence and Signal Processing*. IEEE, 84–88.
[26] Tali Dekel, Michael Rubinstein, Ce Liu, and William T. Freeman. 2017. On the effectiveness of visible watermarks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2146–2154.
[27] Anuja Dixit and Rahul Dixit. 2017. A review on digital image watermarking techniques. *International Journal of Image, Graphics and Signal Processing* 9, 4 (2017), 56.
[28] Yupeng Dong and Chunhui Wang. 2023. Copyright protection on NFT digital works in the Metaverse. *Security and Safety* 2, Article 2023013 (2023), 14 pages.
[29] Dr. Marwan AI Zarouni. 2021. *Dubai's art cars*. https://www.artforcrypto.com/artcar
[30] Haihan Duan, Zhonghao Lin, Xiao Wu, and Wei Cai. 2023. Metacube: A crypto-based unique user-generated content editor for web3 metaverse. *IEEE Communications Magazine* 61, 8 (2023), 52–58.
[31] Haihan Duan, Xiao Wu, and Wei Cai. 2022. Crypto-dropout: To create unique user-generated content using crypto information in metaverse. In *Proceedings of the 2022 IEEE 24th International Workshop on Multimedia Signal Processing*. IEEE, 1–6.
[32] Abdulmotaleb El Saddik, Jamil Ahmad, Mustaqeem Khan, Saad Abouzahir, and Wail Gueaieb. 2025. Unleashing creativity in the metaverse: Generative ai and multimodal content. *ACM Transactions on Multimedia Computing, Communications and Applications* (2025).
[33] Omar Elharrouss, Noor Almaadeed, Somaya Al-Maadeed, and Younes Akbari. 2020. Image inpainting: A review. *Neural Processing Letters* 51 (2020), 2007–2028.
[34] Aviv Gabbay and Yedid Hoshen. 2019. Style generator inversion for image enhancement and animation. arXiv:1906.11880. Retrieved from https://arxiv.org/abs/1906.11880
[35] Emir Ganic, Nasir Zubair, and Ahmet M. Eskicioglu. 2003. An optimal watermarking scheme based on singular value decomposition. In *Proceedings of the IASTED International Conference on Communication, Network, and Information Security*. Citeseer.
[36] Yudong Gao, Xuemei Xie, and Yuan Ni. 2023. Evolutionary game analysis of copyright protection for nft digital works considering collusive behavior. *Applied Sciences* 13, 20 (2023), 11261.
[37] Rafael C. Gonzalez. 2009. *Digital Image Processing*. Pearson education india.
[38] Shanyan Guan, Ying Tai, Bingbing Ni, Feida Zhu, Feiyue Huang, and Xiaokang Yang. 2020. Collaborative learning for faster stylegan embedding. arXiv:2007.01758. Retrieved from https://arxiv.org/abs/2007.01758
[39] Mardi Handono, Ikarini Dani Widiyanti, and Pratiwi Puspitho Andini. 2023. Dispute resolution for non-fungible token (NFT) businesses in Indonesia. *International Journal of Social Science and Education Research Studies* 3, 8 (2023), 1519–1526.
[40] Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross Girshick. 2020. Momentum contrast for unsupervised visual representation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 9729–9738.
[41] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 770–778.
[42] Lital Helman and Ofer Tur-Sinai. 2023. Bracing scarcity: Can NFTs save digital art? *Florida State University Law Review* 51, 1 (2023), 183.
[43] Chun-Hsiang Huang and Ja-Ling Wu. 2004. Attacking visible watermarking schemes. *IEEE Transactions on Multimedia* 6, 1 (2004), 16–30.
[44] Poonam Kadian, Shiafali M. Arora, and Nidhi Arora. 2021. Robust digital watermarking techniques for copyright protection of digital data: A survey. *Wireless Personal Communications* 118 (2021), 3225–3249.
[45] Tero Karras, Miika Aittala, Samuli Laine, Erik Härkönen, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2021. Alias-free generative adversarial networks. *Advances in Neural Information Processing Systems* 34 (2021), 852–863.
[46] Tero Karras, Samuli Laine, and Timo Aila. 2019. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 4401–4410.
[47] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2020. Analyzing and improving the image quality of stylegan. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 8110–8119.
[48] Mustaqeem Khan, Jamil Ahmad, Wail Gueaieb, Giulia De Masi, Fakhri Karray, and Abdulmotaleb El Saddik. 2025. Joint multi-scale multimodal transformer for emotion using consumer devices. *IEEE Transactions on Consumer Electronics* 71, 1 (2025), 1092–1101.
[49] Diederik P. Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. arXiv:1412.6980. Retrieved from https://arxiv.org/abs/1412.6980

[50] Sumit Kumar. 2022. Strategic management of carbon footprint using carbon collectible nonfungible tokens (NFTS) on blockchain. *Academy of Strategic Management Journal* 21 (2022), 1–9.
[51] Apoorv Lal and Fengqi You. 2023. Climate concerns and the future of nonfungible tokens: Leveraging environmental benefits of the Ethereum Merge. *Proceedings of the National Academy of Sciences* 120, 29 (2023), e2303109120.
[52] Michel Legault. 2021. A practitioner's view on distributed storage systems: Overview, challenges and potential solutions. *Technology Innovation Management Review* 11, 6 (2021), 32–41.
[53] Yuqian Lim. 2023. *Centralized vs Decentralized Storage Cost.* Retrieved December 14, 2025 from https://www.coingecko.com/research/publications/centralized-decentralized-storage-cost
[54] Lehao Lin, Hong Kang, Xinyao Sun, and Wei Cai. 2024. SemNFT: A semantically enhanced decentralized middleware for digital asset immortality. In *Proceedings of the 32nd ACM International Conference on Multimedia*. 11051–11059.
[55] Matt Lockyer, Nick Mudge, Jordan Schalm, Sebastian Echeverry, and Zainan Zhou. 2018. *ERC-998: Composable Non-Fungible Token.* Retrieved December 14, 2025 from https://eips.ethereum.org/EIPS/eip-998
[56] Lars Mescheder, Andreas Geiger, and Sebastian Nowozin. 2018. Which training methods for GANs do actually converge?. In *Proceedings of the International Conference on Machine Learning*. PMLR, 3481–3490.
[57] Jarno Mielikainen. 2006. LSB matching revisited. *IEEE Signal Processing Letters* 13, 5 (2006), 285–287.
[58] B. Chandra Mohan and S. Srinivas Kumar. 2008. A robust image watermarking scheme using singular value decomposition. *Journal of Multimedia* 3, 1 (2008), 7–15.
[59] Mohammad Moosazadeh and Gholamhossein Ekbatanifard. 2019. A new DCT-based robust image watermarking method using teaching-learning-based optimization. *Journal of Information Security and Applications* 47 (2019), 28–38.
[60] Mayukh Mukhopadhyay. 2018. *Ethereum Smart Contract Development: Build Blockchain-based Decentralized Applications Using Solidity.* Packt Publishing Ltd.
[61] Long H. Nguyen, Nhat Truong Pham, Mustaqeem Khan, Alice Othmani, and Abdulmotaleb EI Saddik. 2024. HuBERT-CLAP: Contrastive learning-based multimodal emotion recognition using self-alignment approach. In *Proceedings of the 6th ACM International Conference on Multimedia in Asia*. 1–6.
[62] Yotam Nitzan, Amit Bermano, Yangyan Li, and Daniel Cohen-Or. 2020. Disentangling in latent space by harnessing a pretrained generator. arXiv:2005.07728. Retrieved from https://arxiv.org/abs/2005.07728
[63] OpenZeppelin. 2022. *ERC721 Implementation.* Retrieved December 14, 2025 from https://docs.openzeppelin.com/contracts/4.x/erc721
[64] Micah Zoltu Paul Wackerow, Pablo Pettinari. 2025. *Smart Contract on Ethereum.* Retrieved December 14, 2025 from https://ethereum.org/en/developers/docs/smart-contracts/
[65] Soo-Chang Pei and Yi-Chong Zeng. 2006. A novel image recovery algorithm for visible watermarked images. *IEEE Transactions on Information Forensics and Security* 1, 4 (2006), 543–550.
[66] Stanislav Pidhorskyi, Donald A. Adjeroh, and Gianfranco Doretto. 2020. Adversarial latent autoencoders. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 14104–14113.
[67] Immanni Bhanu Prakash, Adarsh Kr Tiwari, and U Hariharan. 2023. Decentralized metadata storage for non-fungible token collections using interplanetary file system. In *Proceedings of the 2023 7th International Conference on Electronics, Materials Engineering and Nano-Technology*. IEEE, 1–6.
[68] National Institute of Standards and Technology. 2012. Secure Hash Standard (SHS). Federal Information Processing Standards Publication 180 - 4. Information Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MD.
[69] Witek Radomski, Andrew Cooke, Philippe Castonguay, James Therien, Eric Binet, and Ronan Sandford. 2018. *ERC-1155: Multi token standard.* Retrieved December 14, 2025 from https://eips.ethereum.org/EIPS/eip-1155
[70] Elad Richardson, Yuval Alaluf, Or Patashnik, Yotam Nitzan, Yaniv Azar, Stav Shapiro, and Daniel Cohen-Or. 2021. Encoding in style: A stylegan encoder for image-to-image translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2287–2296.
[71] Fabian Schär. 2021. Decentralized finance: On blockchain and smart contract-based financial markets. *Review of the Federal Reserve Bank of St Louis* 103, 2 (2021), 153–174.
[72] Johannes Sedlmeir, Reilly Smethurst, Alexander Rieger, and Gilbert Fridgen. 2021. Digital identities and verifiable credentials. *Business and Information Systems Engineering* 63, 5 (2021), 603–613.
[73] P Shalini. 2024. Blockchain technology: Architecture, consensus, and future trends. *Emperor Journal of Applied Scientific Research* 6, 3 (2024), 73–84.
[74] Yujun Shen, Ceyuan Yang, Xiaoou Tang, and Bolei Zhou. 2020. Interfacegan: Interpreting the disentangled face representation learned by gans. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 4 (2020), 2004–2018.
[75] Amritraj Singh, Reza M. Parizi, Qi Zhang, Kim-Kwang Raymond Choo, and Ali Dehghantanha. 2020. Blockchain smart contracts formalization: Approaches and challenges to address vulnerabilities. *Computers and Security* 88 (2020), 101654.
[76] Durgesh Singh and Sanjay K. Singh. 2017. DCT based efficient fragile watermarking scheme for image authentication and restoration. *Multimedia Tools and Applications* 76 (2017), 953–977.

[77] Gilbert Strang. 2000. *Linear Algebra and its Applications*. Pearson Education India.
[78] Solidity Team. 2025. *Solidity Document: Structure of a Contract*. Retrieved December 14, 2025 from https://docs.soliditylang.org/en/latest/structure-of-a-contract.html
[79] Ayush Tewari, Mohamed Elgharib, Florian Bernard, Hans-Peter Seidel, Patrick Pérez, Michael Zollhöfer, and Christian Theobalt. 2020. Pie: Portrait image embedding for semantic control. *ACM Transactions on Graphics* 39, 6 (2020), 1–14.
[80] Jon Toor. 2021. *AWS Prices in Cloudian*. Retrieved December 14, 2025 from https://cloudian.com/blog/5-components-of-aws-s3-storage-pricing/
[81] Omer Tov, Yuval Alaluf, Yotam Nitzan, Or Patashnik, and Daniel Cohen-Or. 2021. Designing an encoder for stylegan image manipulation. *ACM Transactions on Graphics* 40, 4 (2021), 1–14.
[82] Jue Wang and Michael F. Cohen. 2008. Image and video matting: A survey. *Foundations and Trends® in Computer Graphics and Vision* 3.2 (2008), 97–175.
[83] Jinqiao Wang, Qingshan Liu, Lingyu Duan, Hanqing Lu, and Changsheng Xu. 2007. Automatic tv logo detection, tracking and removal in broadcast video. In *Proceedings of the International Conference on Multimedia Modeling*. Springer, 63–72.
[84] Qin Wang, Rujia Li, Qi Wang, and Shiping Chen. 2021. Non-fungible token (NFT): Overview, evaluation, opportunities and challenges. arXiv:2105.07447. Retrieved from https://arxiv.org/abs/2105.07447
[85] Ran-Zan Wang, Chi-Fang Lin, and Ja-Chen Lin. 2001. Image hiding by optimal LSB substitution and genetic algorithm. *Pattern Recognition* 34, 3 (2001), 671–683.
[86] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. 2004. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing* 13, 4 (2004), 600–612.
[87] Ziwei Wang, Jiashi Gao, and Xuetao Wei. 2023. Do NFTs' owners really possess their assets? A first look at the NFT-to-asset connection fragility. In *Proceedings of the ACM Web Conference 2023*. 2099–2109.
[88] William Entriken, Dieter Shirley, Jacob Evans, and Nastassia Sachs. 2018. *ERC721: Non-Fungible Token Standard*. Retrieved December 14, 2025 from https://eips.ethereum.org/EIPS/eip-721
[89] Sam Williams, Viktor Diordiiev, Lev Berman, and Ivan Uemlianin. 2019. Arweave: A protocol for economically sustainable information permanence. *Arweave Yellow Paper* (2019).
[90] Witek Radomski, Andrew Cooke, Philippe Castonguay, James Therien, Eric Binet, and Ronan Sandford. 2018. *ERC-1155: Multi Token Standard*. Retrieved December 14, 2025 from https://eips.ethereum.org/EIPS/eip-1155
[91] Zongze Wu, Dani Lischinski, and Eli Shechtman. 2021. Stylespace analysis: Disentangled controls for stylegan image generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 12863–12872.
[92] Weihao Xia, Yulun Zhang, Yujiu Yang, Jing-Hao Xue, Bolei Zhou, and Ming-Hsuan Yang. 2022. Gan inversion: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 3 (2022), 3121–3138.
[93] Yunpeng Xiao, Bufan Deng, Siqi Chen, Kyrie Zhixuan Zhou, Ray Lc, Luyao Zhang, and Xin Tong. 2024. " Centralized or decentralized?": Concerns and value judgments of stakeholders in the non-fungible tokens (NFTs) market. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–34.
[94] Ceyuan Yang, Yujun Shen, and Bolei Zhou. 2021. Semantic hierarchy emerges in deep generative representations for scene synthesis. *International Journal of Computer Vision* 129 (2021), 1451–1466.
[95] Minqiang Yang, Edith C. H. Ngai, Xiping Hu, Bin Hu, Jiangchuan Liu, Erol Gelenbe, and Victor C. M. Leung. 2025. Digital phenotyping and feature extraction on smartphone data for depression detection. *Proc. IEEE* 112, 12 (2025), 1773–1798.
[96] In-Kwon Yeo and Hyoung Joong Kim. 2003. Generalized patchwork algorithm for image watermarking. *Multimedia Systems* 9 (2003), 261–265.
[97] Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. 2018. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 586–595.
[98] Lei Zhao, Xiaolin Gui, and Yiyang Shao. 2024. Survey on multipurpose digital image watermarking. *Journal of Computer-Aided Design and Computer Graphics* 36, 2 (2024), 195–222.
[99] Jiapeng Zhu, Yujun Shen, Deli Zhao, and Bolei Zhou. 2020. In-domain gan inversion for real image editing. In *Proceedings of the European Conference on Computer Vision*. Springer, 592–608.
[100] Jun-Yan Zhu, Philipp Krähenbühl, Eli Shechtman, and Alexei A. Efros. 2016. Generative visual manipulation on the natural image manifold. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, the Netherlands, October 11-14, 2016, Proceedings, Part v 14*. Springer, 597–613.

Received 1 April 2025; revised 17 August 2025; accepted 19 November 2025